EXPLORING GENE EXPRESSION STRATEGIES IN POTATO LEAFROLL VIRUS

JULIA ARRIGO

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIRMENTS FOR THE DEGREE OF MASTER OF SCIENCE

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY TORONTO,

ONTARIO CANADA

January 2024

ABSTRACT

Potato leafroll virus (PLRV) is a positive-sense single-stranded RNA plant virus in the genus *Polerovirus* (family *Solemoviridae*). This virus is known to use many gene expression strategies during viral infection, including subgenomic messenger RNA (sg mRNA) transcription and non-AUG initiation. In addition, it is known to produce a small viral RNA (svRNA) degradation product by cellular exoribonuclease digestion. The goal of this thesis was to investigate PLRV gene expression strategies and corresponding translational products. Two conserved structural elements, a sg mRNA promoter element, and a downstream stem loop (dSL), were identified. A previously defined RNA element involved in generating svRNA was also investigated. The results uncovered a sg mRNA transcriptional promoter element, a potential coding function for svRNA and no role for the dSL in modulating viral protein translation. These findings contribute to understanding how PLRV expresses its viral proteins during infections.

ACKNOWLEDGEMENTS

First, I would like to extend my deepest appreciation for my supervisor, Dr. K. Andrew White. I would not have been a successful graduate student without his support, encouragement, and guidance.

I am extremely grateful for my committee member Dr. Katalin Hudak, whose comments, suggestions and evaluations were very beneficial for the progression of my research.

I'd like to extend my gratitude to current and previous White lab members. I'm grateful for Dr. Tamari Chkuaseli, I appreciate her patience while mentoring and teaching me valuable lab techniques. Her advice, friendship and belief in my abilities kept me motivated throughout my Masters. I'd like to thank Dr. Baodong Wu for providing lab materials and helpful advice when needed, and Dr. Chaminda Gunawardene and Jennifer Im for their valuable friendship and support.

I'd also like to acknowledge assistance from Hudak and Bayfield labs for providing me with additional lab materials, when needed.

Lastly, I would like to thank my family and friends for their love, encouragement, and support throughout my studies.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ABBREVIATIONS

(+) ssRNA: positive-sense single strand RNA

3′-CITE: 3′-Cap Independent Translation Enhancer

A: adenine

aatRNA: amino acyl tRNAs

AGO1: ARGONATE1

AMP: ampicillin

APS: Ammonium persulfate

ARCA: anti-reverse cap analog

ATP: Adenosine triphosphate

BME: 2-Mercaptoethanol

BMV: Bromo mosaic virus

bp: base pairs

BrYV: Brassica yellows virus

BSA: Bovine serum albumin

BWYV: Beet western yellows virus

BYDV: Barley Yellow Dwarf virus

C-: carboxyl

cDNA: copy DNA

CIP: Calf-intestinal Alkaline Phosphatase

CP: coat protein

CPPV1: Cowpea polerovirus 1

CP-RTD: coat protein-readthrough domain

CTP: cytosine triphosphate

CtRLV: Carrot red leaf virus

CYDV: Cereal yellow dwarf virus

ddH_2O: double distilled water

DEPC H_2O: diethyl pyrocarbonate-treated water

DNase: deoxyribonuclease

dNTP: deoxyribonucleotides triphosphate

DRTE: distal readthrough element

dSL: downstream stem loop

DTT: dithiothreitol

E. coli: Escherichia coli

EDTA: ethylenediaminetetraacetic acid

eIF: eukaryotic initiation factor

ER: endoplasmic reticulum

eRF: eukaryotic release factor

EtBr: ethidium bromide

G: guanine

$\gamma^{32}P$: gamma 32-phosphate

GMP: guanosine monophosphate

GTP: guanosine triphosphate

HA: hemagglutinin

IRES: internal ribosomal entry site

kDa: kilo Dalton

KO: knock out

M: Molarity

M^7G cap: 7-methylguanosine cap

mA: milliamps

Met: methionine

Met-$tRNA_i$: initiator methionyl-tRNA

µg: microgram

µl: microliter

µM: micromolar

mM: millimolar

MP: movement protein

mRNA: messenger RNA

NEB: New England Biolabs

nt: nucleotide

NTP: nucleoside triphosphate

OD: optical density

-OH: hydroxyl

OPMV: Opium poppy mosaic virus

ORF: open reading frames

PABP: poly (A) binding protein

PCI: phenol chloroform isoamyl-alcohol

PCR: polymerase chain reaction

PEG: polyethylene glycol

PEMV: Pea Enation Mosaic virus

PeVYV: Pepper vein yellows virus

PIC: preinitiation complex

pk: pseudoknot

PLRV: Potato leafroll virus (PLRV)

pmols: picomoles

PNK: polynucleotide

poly (A) tail: poly adenosine tail

PRTE: proximal readthrough element

P-site: peptidyl site

Rap1: replication associated protein 1

Rcf: relative centrifugal field

RCNMV: Red Clover Necrotic Mosaic virus

RdRp: RNA dependent RNA polymerase

RISC: RNA induced silencing complex

RNase: ribonuclease

rpm: revolutions per minute

RSV: Rous sarcoma virus

S^{35}-Met: 35-Sulfur methionine

ScYLV: Sugarcane yellow leaf virus

SDS: sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

sg: subgenomic

SgP: sg mRNA promoter

sgRNAs: sg mRNA + svRNA band on northern blot

SHAPE: Selective 2'-Hydroxyl Acylation Analyzed by Primer Extension

SLB: super Luria-Bertani broth

SPV1: Strawberry polerovirus 1

SSC: saline-sodium citrate

STE: sodium chloride tris EDTA

sTNV: Satellite tobacco necrosis virus

sv: small viral

TAE: Tris-acetate-EDTA

TBE: Tris Boric acid EDTA

TBSV: Tomato bushy stunt virus

TC: ternary complex

TCV: Turnip crinkle virus

TEMED: tetramethyl ethylenediamine

TMV: Tobacco mosaic virus

TNV-D: Tobacco necrosis virus-D

tRNA: transfer RNA

TRS: transcription regulatory sequence

TVDV: Tobacco vein distorting virus

U: uracil

UTP: uridine triphosphate

UTR: untranslated region

V: volts

VPg: viral protein genome-linked

Wge: wheat germ extract

WYDV: Wheat yellow dwarf virus

Xrn: exoribonuclease

xrRNA: exoribonuclease resistant RNA structures

1 Introduction

1.1 Positive-sensed single stranded RNA plant viruses

Viruses are known as obligate intracellular parasites that could be described simply as genes packaged inside a surrounding protein coat (Koonin et al., 2020). The viral genome is conventionally safely enclosed into a protein capsid, and together the genome and capsid are known as a virion (**shown in Figure 1 viral life cycle**). Capsids take on various symmetries in the form of icosahedral or helical protein shells, which help to protect the genome from harsh environments before it is delivered to a host cell to initiate the viral reproductive cycle (Sevvana et al., 2021). Viral genomes are classified into several different classes based on nucleic acid type (Koonin et al., 2021). One classification is positive-sense single-stranded RNA [(+) ssRNA] viruses, this is the most common type of plant virus, which are known to cause loss and damage to economically important crops worldwide (Koonin et al., 2020; Newburn & White, 2015). Positive-sensed RNA virus genomes are single-stranded, coding-sensed, RNA molecules, therefore once a genome enters its host cell it is directly translatable using cellular translation machinery (Koonin et al., 2021; Miras et al., 2017).

There are distinct steps that are involved in the viral reproduction cycle, as shown in **Figure 1**. The first step is **1)** entry of the viral RNA into the cytosol of the host cell. In plant cells, virions must penetrate the host cell wall through breaks or injuries in the plant, usually caused by biological damage from feeding insect vectors (Dimitrov, 2004; Kozieł et al., 2021). After penetration, the virion must **2)** disassemble or uncoat, to release its RNA genome. The mechanism(s) are mostly unknown, but in tobacco mosaic virus co-translational and co-replicational strategies may be involved (Newburn & White, 2015). The next step is **3)** translation of the positive-strand viral genome by host ribosomes. This generates replication proteins such as viral RNA-dependent RNA polymerase (RdRp) responsible for genome replication and subgenomic (sg) mRNA transcription (Jiwan & White, 2011; Newburn & White, 2015). The expression of viral proteins takes on many unconventional translational strategies, such as cap-independent translation or recoding strategies, which are described in later sections (Firth & Brierley, 2012; Geng et al., 2021; Miras et al., 2017). The RdRp, along with the help of host proteins, are used for **4)** replication of the viral genome, via a negative-strand RNA

intermediate, and the progeny generated then serve as templates for further translation or sg mRNA transcription. Both genome replication and sg mRNA transcription occur in membrane enclosures in order to avoid detection of double-stranded RNA intermediates by the plant antiviral RNA interference immune system (Nagy et al., 2016; Rampersad & Tennant, 2018).

Many plus-strand RNA viral genomes encode proteins that are located far from the 5'-end. Due to the 5'-end dependence of the ribosome, transcription of smaller viral messages called sg mRNAs occur by **5)** transcription performed by the viral RdRp during later steps of infections. These sg mRNAs are 3'-coterminal with the viral genome, but have 5'-ends corresponding to 3'-proximally positioned proteins, thus allowing for **6)** translation of "late" genes such as coat protein (CP) or movement protein (MP) (Jiwan & White, 2011; Newburn & White, 2020; White, 2002). The final intracellular step of the viral reproductive cycle is **7)** virion assembly, also known as packaging, where CP subunits encase the viral genome. The infection can then spread to adjacent cells through plasmodesmata and systemically through the plant circulatory system, and finally to other plants by insect vectors (Kumar et al., 2015; Shi et al., 2021). Although not shown in **Figure 1**, some viruses also produce small RNA degradation products via cellular exoribonucleases, herein referred to as small viral (sv) RNAs. Some of these svRNAs encode viral proteins and therefore can act as viral mRNAs (Gunawardene et al., 2019; Nagarajan et al., 2013; Steckelberg et al., 2018).

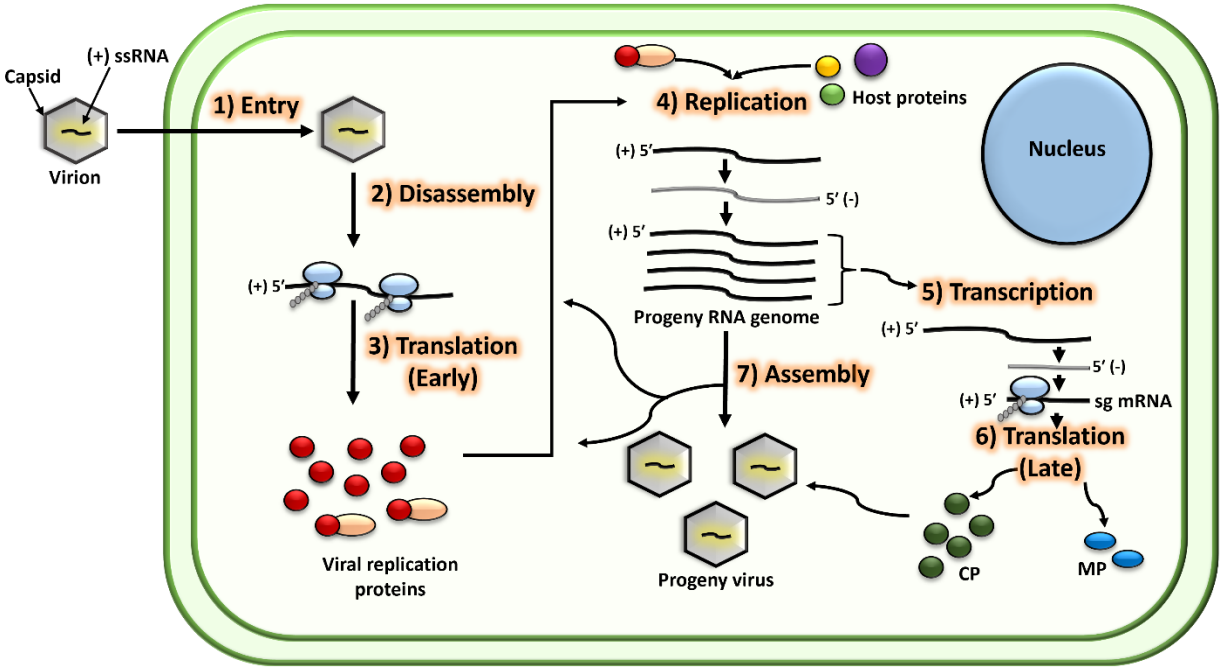


Figure 1. Simplified step by step outline of a positive-sense RNA plant viral reproductive cycle. Each step of the cycle is numbered, and labels are highlighted in orange. **1)** The virion (grey hexagon depicts the capsid, while the black small squiggly line highlighted in yellow depicts the positive-sensed single stranded RNA [(+) ssRNA)] first must enter the plant cell through breaks in the cell wall. **2)** Once inside the plant cell, the viral particle disassembles exposing its positive-sense RNA genome for translation. **3)** Early translation involves using the plus-strand RNA genome for translation of viral replication proteins (red circle and orange oval combined makes up the viral replication protein) that are then used for step **4)** viral genome replication. During this step, the replication proteins, along with host factors (purple, yellow and green circles), are employed to generate a minus-strand RNA (grey squiggly line) that is used as a template to generate more full-length plus-strand RNA. These plus-strand RNAs could be packaged into viral particles, act as a template for early translation or be used for step **5)** transcription. Transcription uses the replication proteins to transcribe subgenomic messenger RNAs (sg mRNA) that encode proteins located 3′-coterminally in the genome. The sg mRNAs are then used as a template for **6)** late translation to generate proteins such as coat protein (CP, shown in green) or movement protein (MP, shown in blue) that are required later in the infection cycle. Finally, the CP generated during late translation are used for **7)** assembly of the viral particle (Figure was adapted from White, 2011).

1.2 Translation initiation strategies used by positive-sensed single-stranded RNA plant viruses

Most eukaryotic mRNAs have a 5'-cap and 3'-poly (A) tail, that are crucial structures involved in mediating canonical translation. Translation consists of four steps initiation, elongation, termination, and ribosomal recycling (Firth & Brierley, 2012).

Translation initiation in eukaryotic cells involves: formation of the 43S ribosomal subunit preinitiation complex (PIC) and recognition of the messenger RNA (mRNA), scanning, start codon recognition, and 80S ribosome formation (Aitken & Lorsch, 2012). First, the 43S PIC ribosome subunit is formed, that includes the ternary complex (TC), eIF2-Met-tRNA-GTP. Before this complex is recruited to the mRNA, the eukaryotic initiation factor (eIF) 4E (eIF4E) recognizes the 5'-cap structure, in association with eIF4G (Aitken & Lorsch, 2012; Hoang et al., 2021). As a scaffold protein, the task of eIF4G is to recruit helicase eIF4A and eIF4B. The poly (A) binding protein (PABP) is bound to the poly (A) tail at the 3'-end of the mRNA. PABP interacts with eIF4G at the 5'-end leading to mRNA circularization (Firth & Brierley, 2012; Wells et al., 1998). Once the 5'- and 3'-end are together, the PIC is efficiently recruited to the mRNA, followed by ribosomal scanning. The 43S ribosomal subunit scans the mRNA in search of an AUG start codon, that is in an optimal "Kozak" context. For eukaryotic mRNA an AUG in a strong initiation context has G at the +4 position and a purine in the -3 position (Firth & Brierley, 2012; Kozak, 1981). Once the ribosomal subunit finds the initiation codon a Met-tRNA$_i$ is placed in the P-site of the ribosome. Finally, addition of the 60S ribosomal subunit, allows the formation of the 80S ribosome (Firth & Brierley, 2012). After this final step, translation proceeds to elongation, which involves a continuous supply of amino-acyl tRNAs (aatRNAs) used for peptide bond formation and ribosomal translocation in the 3'-direction. Once a stop codon is reached eukaryotic release factors (eRFs) are recruited to the ribosome, resulting in polypeptide chain release. This is followed by ribosome recycling, which involves dissociation of the 80S ribosome (Dever & Green, 2012; Hoang et al., 2021).

Unlike eukaryotic mRNA, many positive-strand plant RNA viruses do not have a 5'-cap or 3'-poly A tail, therefore in order to compete with host ribosomes alternative strategies are used. These can involve RNA elements located at the 5'-end, 3'-end, or internally within the viral genome. These translational RNA elements are described in the next section.

1.2.2 Cap independent translation strategies

1.2.2.1 3’-Cap Independent Translation Enhancers (3’-CITEs)

As the name suggests, a 3’ cap-independent translation enhancer (3’-CITE) is a type of cap-independent translation strategy, and is commonly found in (+) ssRNA plant viruses (Geng et al., 2021). These mRNA elements are known to recruit translation initiation factors, such as eIF4G and eIF4E, to the 3’-end of viral mRNAs (Newburn & White, 2015; Nicholson et al., 2010). Due to translation being initiated at the 5’-end of the mRNA, in many cases there is a 5’-untranslated region (UTR) to 3’-CITE long-distance base pairing interaction (**Figure 2A**) that delivers initiation factors bound to the 3’-CITE to the 5’-end of the genome, which then recruits the 43S ribosomal subunit for scanning of the viral RNA (Nicholson & White, 2011).

There are various structures of 3’-CITE’s, the first one was discovered in satellite tobacco necrosis virus (sTNV), a type species of the *Tombusviridae* family. This 3’-CITE consisted of 619 nts located in the 3’-UTR. It was originally thought to initiate coat protein expression through long distance interaction with the 5’-UTR, however it was later shown that base pairing between the two ends is not responsible for its interaction (Danthinne et al., 1993; Geng et al., 2021; Meulewaeter et al., 1998). A 3’-CITE that involves 5’-to-3’ base pairing is found in barley yellow dwarf virus (BYDV). This structure includes a highly conserved 17 nts sequence and stable stem loop that is involved in pairing with the 5’-UTR (Geng et al., 2021; Wang et al., 2010). This 3’-CITE has a very high affinity for binding to eIF4G, which in association with eIF4E recruits translational machinery for translation initiation (Treder et al., 2008).

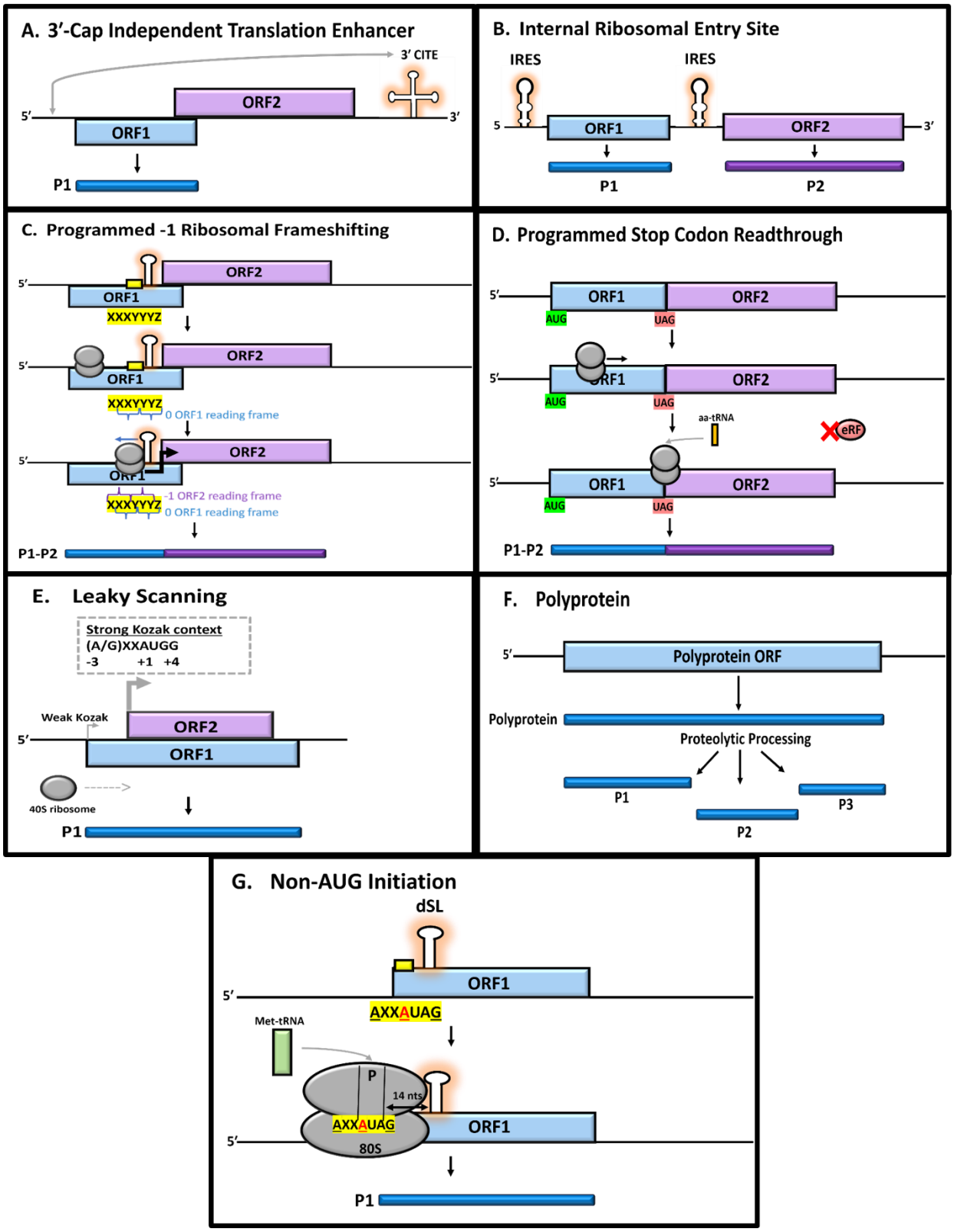

A. 3′-Cap Independent Translation Enhancer
3′ CITE
5′
ORF2
ORF1
3′
P1
B. Internal Ribosomal Entry Site
IRES
IRES
5
ORF1
ORF2
3′
P1
P2
C. Programmed -1 Ribosomal Frameshifting
5′
ORF2
ORF1
XXXYYYZ
0 ORF1 reading frame
-1 ORF2 reading frame
P1-P2
D. Programmed Stop Codon Readthrough
AUG
UAG
aa-tRNA
eRF
P1-P2
E. Leaky Scanning
Strong Kozak context
(A/G)XXAUGG
-3 +1 +4
Weak Kozak
40S ribosome
P1
F. Polyprotein
Polyprotein ORF
Polyprotein
Proteolytic Processing
P1
P2
P3
G. Non-AUG Initiation
dSL
AXXAUAG
Met-tRNA
P
14 nts
80S
P1

Figure 2. Non-canonical translation initiation strategies used by positive-sense RNA plant viruses. Plus-strand viral genomes are shown with blue and purple boxes representing different open reading frames (ORFs). The blue or purple bars depict corresponding protein products. **A.** A 3'-cap independent translation enhancer (3'-CITE) is the black stem loop structure at the 3'-end that is highlighted in orange. It recruits translation initiation factors to the 3'-end of the virus and relocates them to the 5'-end via base pairing interactions with the 3'-CITE (depicted with grey arrow). **B.** Internal ribosomal entry site (IRES) located at the 5'-end or internally within the viral genome (represented by black stem loop structures highlighted in orange) they function to recruit ribosomes. **C.** Programmed -1 ribosomal frameshift involves a heptanucleotide sequence depicted by yellow highlighted nucleotides (nucleotides represented by each letter are shown in section 1.2.3.1) in the blue ORF1 and a stem loop highlighted in orange in ORF1. A ribosome (depicted in grey) translating ORF1 will pause on the heptanucleotide sequence due to the downstream structure and slip back one nucleotide, represented by blue arrow shown pointing in the opposite direction of the translating ribosome. This results in switching reading frames from 0 (shown in blue depicting ORF1 reading frame) to -1 (shown in purple depicting ORF2 reading frame) leading to a protein product consisting of ORF1 and ORF2 together (P1-P2). **D.** Programmed stop codon readthrough, involves a translating ribosome misreading a stop codon (highlighted in red) as a sense codon resulting in failed termination (depicted by red X in front of red eukaryotic release factor (eRF)). Leading to continued translation (depicted by the addition of an orange amino acyl t-RNA (aatRNA)) into the purple ORF2 that is in the same reading frame as ORF1, producing a P1, C-terminal extended product, P1-P2. **E.** Leaky scanning. For translation initiation, a scanning 40S ribosomal subunit (depicted by grey circle) scans pass a weak Kozak context, shown by the small grey arrow at the 5'-end of the blue ORF1, and contiunes scanning until it reaches a start codon in a strong Kozak context (depicted in box outline by dashed line and by the thick grey arrow at the 5'-end of the purple ORF2). **F.** Polyprotein processing involves proteolytic cleavage of a long polypeptide chain (depicted by dark blue long thin rectangle) into multiple proteins (smaller rectangles labeled as P1, P2 and P3). **G.** Non-AUG initiation involves the use of a non-AUG start codon in a strong Kozak context shown highlighted in yellow and may involve a downstream stem loop (dSL), highlighted in orange. The dSL acts to pause the ribosome approximately 14 nts away from the start codon, which promotes alignment of the start codon with the P-site of the ribosome (depicted by Met-tRNA green and light grey arrow pointing towards P-site of ribosome). This results in translation of ORF1 and production of protein product P1 (dark blue thin rectangle).

1.2.2.2 Internal Ribosomal Entry Sites (IRESes)

Internal ribosomal entry sites (IRESes) can involve very complex *cis*-acting RNA structures located upstream of an initiating AUG, either at the 5'-end of the viral genome or internally within the genome (**Figure 2B**). Its function is to recruit ribosomes. The discovery of the first IRES was made in the family *Picornaviridae*, specifically encephalomyocarditis and poliovirus. These viruses had IRESes that were found to be very long RNA sequences of at least 450 nts in length (Geng et al., 2021; Pelletier & Sonenberg, 1988). In plant RNA viruses, IRESes are shorter RNA sequences starting from 60-190 nts that are not as complex as animal virus IRESes. In potyviruses, IRESes are pseudoknot structures, located in the 5'-UTR and function by binding to eIF4G (Geng et al., 2021; Newburn & White, 2015; Zeenko & Gallie, 2005). In plant viruses, IRESes can also be present internally within the viral genome and can be involved in the translation of proteins such as coat protein (Dorokhov et al., 2006; Newburn & White, 2015). A very unusually IRES is found in the family of *Dicistroviridae* that is 180 nts long and is folded using pseudoknots forming a structure that mimics tRNAs. This structure acts to initiate translation by allowing the formation of the 80S ribosome complex required to progress in the translation cycle (Firth & Brierley, 2012; Jang et al., 2009).

1.2.3 Translation recoding strategies

Due to the small and limiting size of viral RNA genomes many are composed of overlapping open reading frames (ORF), and this is beneficial as it maximizes their coding capacity. The disadvantage to overlapping ORFs are that not all genes are readily accessible to host translational machinery, so viruses have adapted alternative translational strategies to expression these overlapping genes, such as recoding translation strategies (programmed -1 ribosomal frameshift, programmed stop codon readthrough) and leaky scanning (Firth & Brierley, 2012; Geng et al., 2021; Miras et al., 2017).

1.2.3.1 Programmed -1 Ribosomal Frameshifting

During this translation strategy the ribosome encounters two cis-acting elements that are separated by a spacer sequence. There is a heptanucleotide slippery site (**Figure 2C**) on which the ribosome slips one nucleotide in the 5' direction into the -1 reading frame (**Figure 2C**). This sequence is followed by a downstream RNA secondary structure (**Figure 2C**), which

pauses the ribosome that facilitates its backstepping (Riegger & Caliskan, 2022). This results in the ribosome translating two genes encoded in different reading frames. The *cis*-elements involved in -1 frameshifting are the heptanucleotide slippery site, which is composed of nucleotides XXXYYYZ (where X is any base, Y is an adenine (A) or uracil (U), Z is any nucleotide but guanine (G)) and a downstream stimulatory RNA structure that can be a stem loop or pseudoknot (Firth & Brierley, 2012; Jacks et al., 1988; Riegger & Caliskan, 2022).

This strategy was first found to be involved in Rous sarcoma virus (RSV) for the expression of the Gag-Pol polyprotein. Pol is responsible for reverse transcription, while Gag is involved in packaging and forming the capsid. The mechanism involved using a slippery sequence as well as a RNA stem loop in order to express the overlapping Gag and Pol ORFs. Due to the protein's individual functions, the Gag-Pol product was suggested to be a way of regulating replication and packaging for RSV (Geng et al., 2021; Jacks et al., 1988; Riegger & Caliskan, 2022). This mechanism of translation is also found in plant RNA viruses such as the polerovirus, potato leafroll virus (PLRV). In this virus the mechanism is used to fuse two proteins P1 and P2 that have overlapping reading frames (**Figure 5**). The P1-P2 fusion protein is known as the RdRp involved in viral replication (Delfosse et al., 2021). This mechanism will be described in more detail in the PLRV section. Other plant RNA viruses such as pea enation mosaic virus 2 (PEMV2) in the *Tombusviridae* family also uses -1 ribosomal frameshifting for the production of its RdRp (Gao & Simon, 2016).

1.2.3.2 Programmed Stop Codon Readthrough

The process of programmed stop codon readthrough involves a ribosome misreading a stop codon as a sense codon, resulting in failed termination of translation (**Figure 2D**) (Geng et al., 2021). Instead, a near-cognate tRNA is used and an amino acid is inserted in the polypeptide allowing translation to continue in the same reading frame (**Figure 2D**). This differs from a frameshifting event because the ribosome is allowed to continue translating in the original reading frame, not an alternate reading frame (Miras et al., 2017; Rodnina et al., 2020).

The largest group of viruses that use readthrough are positive-sense RNA viruses such as tombusviruses, luteoviruses, poleroviruses and enamoviruses. Tombusviruses use this strategy to express their RdRp for replication, while poleroviruses or enamoviruses translate a C-

terminally extended minor capsid protein using readthrough (Chkuaseli & White, 2022; Newburn et al., 2014). Some programmed stop codon readthrough processes involve specific codons that follow the stop codon, along with a stem loop, while others involve long-distance RNA-RNA interactions such as the betanecrovirus, tobacco necrosis virus-D (TNV-D; Newburn & White, 2015). In this latter virus, readthrough is used to generate the RdRp. The readthrough signal for this virus involves elements known as proximal readthrough element (PRTE), located 3' to the readthrough site and distal readthrough element (DRTE) located near the 3'-end of the virus (Newburn et al., 2014). Compensatory mutational analysis showed that base pairing between both elements, along with the stop codon identity (UAA or UGA), are required for efficient readthrough translation of the RdRp (Newburn et al., 2014). The long-range base pairing interaction has been shown to regulate translation and replication processes in type viral species of the *Tombusvirus* genus. For example, the interaction of the PRTE-DRTE elements inhibit the activity of replication elements. This allows translational readthrough to occur while inhibiting minus-strand synthesis, which would interfere with translation (Cimino et al., 2011).

In the *Polerovirus* and *Enamovirus* genera, programmed stop codon readthrough is involved in creating a C-terminally extended minor coat protein termed CP-RTD (coat protein-readthrough domain) that facilitates systemic movement throughout the plant and mediates transmission of the virus from plant to plant by aphids (Bruyère et al., 1997; Delfosse et al., 2021; Peter et al., 2008). Readthrough in both viral genera involve a complex sequence of interactions between the PRTE-DRTE elements located slightly near the 3'-end of the CP ORF and near the 3'-end of the CP-RTD ORF, respectively. These interactions are responsible for assembly of the readthrough structure for efficient translation of CP-RTD (Chkuaseli & White, 2022; Xu et al., 2018).

1.2.4 Leaky Scanning

Leaky scanning is described as a translational strategy that occurs when a ribosome bypasses a start codon due to its weak or non-optimal Kozak sequence, causing a failure in translation initiation (**Figure 2E**). Instead, the ribosome continues scanning downstream until it reaches a second more optimal translational start site, *e.g.* with A/G at position -3 and G at position +4 (**Figure 2E)** (Firth & Brierley, 2012; Geng et al., 2021; Kozak, 1986; Ryabova et al., 2006). This translation strategy was first discovered in the genus *Orthobunyavirus*. In this virus

two ORFs are located in overlapping reading frames, with one smaller ORF being completely engulfed by the other larger ORF (Fuller et al., 1983). Another example of leaky scanning is in a virus from the *Luteoviridae* family for translation of its movement protein (ORF4). Similar to PLRV, some ribosomes fail to initiate translation on the CP start codon and therefore start translation on the MP start codon (ORF4) (Dinesh-Kumar & Miller, 1993). Leaky scanning mechanism could also occur with ORFs that have non-AUG start codons, such as AUA or CUG (Miras et al., 2017). Further description on non-AUG initiation sites will be described in section 1.2.6.

1.2.5 Polyprotein processing

A strategy used to express ORFs that are not found near the 5'-end of the viral RNA genome is proteolytic processing of long polypeptide chains into multiple proteins (**Figure 2F**). This occurs in the animal *Picornaviridae* family (*e.g.* poliovirus) that produce a polyprotein with the molecular weight of approximately 250 kDa (Spall et al., 1997). Polyprotein processing strategy has its advantages and disadvantages. Some advantages include overcoming the problem of start codons of individual proteins not being accessible and the production of many functional proteins from a limited amount of genetic information. Some disadvantages include less control of protein levels, for example, there must be an abundance of CPs made to encapsidate viral genomes, whereas only one replication protein is needed for replication. This is also a problem in plant potyviruses (Hull, 2001; Miras et al., 2017). To solve this issue some viral groups have divided their genetic information between two or more nucleic acid segments, or they use alternative translation strategies such as those described above (Spall et al., 1997).

1.2.6 Non-AUG initiation translation

Some viruses have ORFs that initiate with non-AUG start codons (**Figure 2G**), such as CUG, GUG, ACG, AUU, AUA, AUC and UUG (Firth & Brierley, 2012; Geng et al., 2021). These non-AUG start codons are generally recognized inefficiently and require the presence of a strong Kozak sequence in order to start translation (Kozak, 1990). An optimal Kozak sequence can also act along with a downstream RNA stem loop (dSL) structure, present about 14 nts away from the non-AUG start codon (**Figure 2G**). The dSL causes pausing of the ribosome on the non-AUG start codon and facilitates its recognition (**Figure 2G**; Kozak, 1990). In viruses, this mechanism is used for the expression of multiple isoforms of the same protein, or in the expression of distinct

proteins (Touriol et al., 2003; Turina et al., 2000). For example, in the tombusvirus, panicum mosaic virus, the first movement protein has an AUG initiation codon, while the second movement protein is expressed using a non-AUG initiation codon, GUG, possibly involving a dSL (Turina et al., 2000).

1.3 Subgenomic messenger RNA (sg mRNA) transcription

Another strategy used by viruses, due to 5'-end dependence for translation, is the transcription of sg mRNAs. Sg mRNAs consist of viral coding sequences and are used for the translation of different viral proteins, such as CP or MP (Firth & Brierley, 2012; Kozak, 1989). These sg mRNAs are transcribed by the viral polymerase and result in the relocation of ORFs at the 3'-end of the viral genome to the 5'-end of a sg mRNA, resulting in efficient translation by the host machinery (Jiwan & White, 2011; White, 2002). There are three viral sg mRNA transcription mechanisms currently known, internal initiation, discontinuous synthesis, and premature termination (**Figure 3**) (Sztuba-Solińska et al., 2011), described in the sections below.

1.3.1 Internal Initiation transcription mechanism

This mechanism of transcription involves the RdRp using the (+)-strand viral genome to generate a full-length (-)-strand of the genome, which is then used as a template for sg mRNA transcription (**Figure 3A**). The full-length (-)-strand genome has a sg mRNA transcriptional promoter located internally (Koev & Miller, 2000; Sztuba-Solińska et al., 2011). This mechanism was first seen in brome mosaic virus (BMV). This virus is composed of three segments and the third segment, RNA3, uses an internal initiation mechanism for production of sg RNA4. The BMV sg promoter element includes an A/U rich enhancer, a core element, and a hairpin, and all elements are involved in binding the RdRp (French & Ahlquist, 1988; P. C. Haasnoot et al., 2000; Siegel et al., 1997). This mechanism is also employed by monopartite RNA viruses such as tobacco mosaic virus (TMV). TMV sg mRNA promoter element consists of a core region and RNA hairpins like BMV, however it was also found to involve long distance interactions with regions in the 3'-UTR to enhance sg mRNA production (Shivprasad et al., 1999; Sztuba-Solińska et al., 2011).

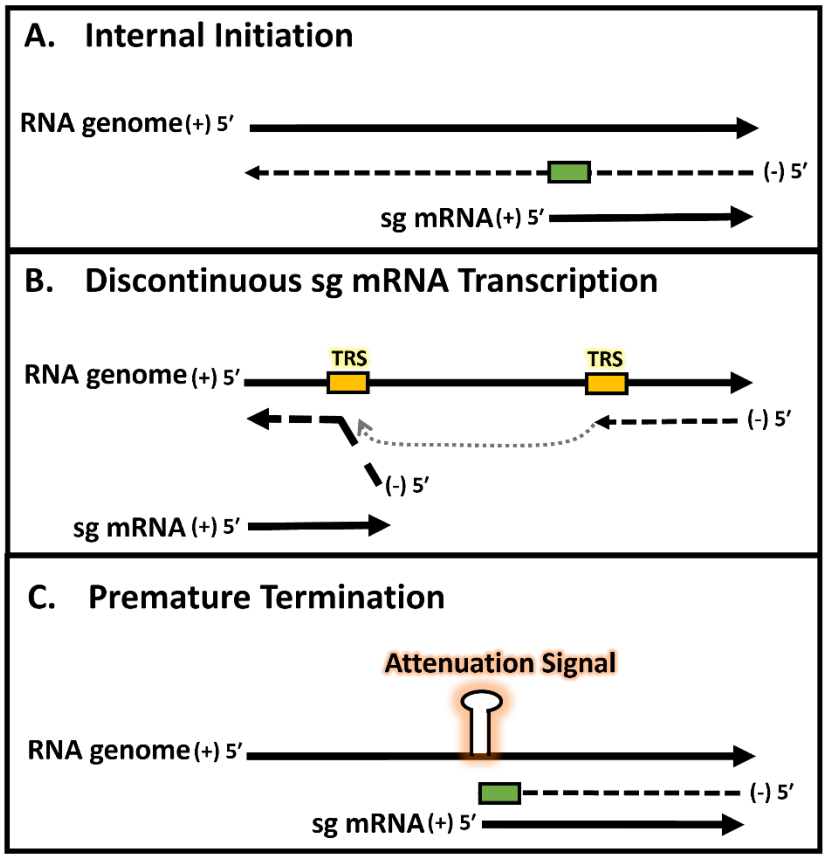


Figure 3. Subgenomic messenger RNA (sg mRNA) transcriptional mechanisms. Solid black arrows depict plus-strand full-length RNA genomes and sg mRNA. Dashed black arrows depict minus-strand full-length RNA genomes and sg mRNA. **A.** Internal initiation involves the RNA dependent RNA polymerase (RdRp) using the full-length plus-strand RNA genome to transcribe a full-length minus-strand copy of the genome. An internal promoter (green box) in the minus-strand is used to generate the sg mRNA. **B.** Discontinuous sg mRNA transcription uses transcription regulatory sequences (TRS, yellow boxes on the RNA genome) to facilitate translocation of the RdRp from one TRS to the other. This generates a minus-strand copy of the sg mRNA that is then used as a template to transcribe the sg mRNA. **C.** Premature termination mechanism employs an attenuation signal (highlighted in orange) that prevents the RdRp from continuing transcription. This results in the formation of a truncated minus-strand that is used as a template to transcribe the plus-strand sg mRNA (Figure adapted from White, 2002).

1.3.2 Discontinuous transcription of sg mRNAs

In the discontinuous synthesis transcription mechanism, the (+)-strand RNA genome is used as a template to synthesize a non-contiguous (-)-strand RNA, that is then copied by the RdRp to produce a sg mRNA (**Figure 3B**). This mechanism of transcription involves elements such as the core sequence, and transcription regulatory sequence (TRS) (Sztuba-Solińska et al., 2011). Steps such as transcription initiation complex formation, scanning for complementary base pairing interaction and template switching are involved in the creation of a nested set of sg mRNAs (Acheson, 2011; Sola et al., 2015). Discontinuous transcription in coronaviruses starts with the RdRp initiating (-)-strand synthesis at the 3'-end of the (+)-strand genome, and it continues to copy until it reaches a TRS. Once it reaches and copies the TRS it will dissociate from the (+)-strand RNA, in search for a complementary upstream TRS sequence located in the 5'-UTR region of the viral genome, where it will reprime and continue synthesis (**Figure 3B, TRS highlighted in yellow, corresponds to yellow box on the RNA genome**). Complementary base pairing allows the 3'-end of the nascent (-)-strand RNA to serve as a primer at the upstream TRS. This forms a "spliced" (-)-strand sg mRNA template that is used by the RdRp to transcribe (+)-strand sg mRNA (Acheson, 2011; Sola et al., 2015).

1.3.3 Premature termination mechanism

For this mechanism of sg mRNA transcription, the RdRp starts copying the (+)-strand RNA genome, starting from its 3'-end. However, instead of creating a full length (-)-strand, it prematurely terminates synthesis, generating a truncated sg mRNA-sized (-)-strand (**Figure 3C**). This small (-)-strand template is then used for transcription of sg mRNA, which can then be translated (Jiwan & White, 2011; White, 2002). Synthesis of sg mRNA via the premature termination mechanism requires an attenuation signal present as an RNA secondary or tertiary structure and created through *cis* or *trans* RNA-RNA interactions (White, 2002). This type of transcriptional mechanism is most commonly found in viruses from the *Tombusviridae* family (Jiwan & White, 2011; Sztuba-Solińska et al., 2011; White, 2002). For tombusviruses, a series of long-distance RNA-RNA interactions in *cis* are required as an attenuation signal for production of sg mRNAs (White & Nagy, 2004). A dianthovirus, red clover necrotic mosaic virus (RCNMV), also produces an sg mRNA through premature termination mechanism. However, this virus is bi-

segmented and uses an attenuation signal composed of a *trans*-base pairing interaction between sequences found in RNA1 and RNA2 (Sit et al., 1998).

1.4 Eukaryotic exoribonuclease (Xrn) degradation

RNA decay pathways for recycling nucleotides are important for cellular gene expression control and occurs by two major pathways 3'-to-5' decay or 5'-to-3' decay (Mugridge et al., 2018). In 3'-to-5' decay an RNA exosome is used to degrade transcripts from their 3'-ends. In 5'-to-3' decay, an exoribonuclease is responsible for degrading transcripts from their 5'-ends (**Figure 4**). For cellular transcripts this multistep process of 5'-to-3' RNA decay involves deadenylation (shorting of the poly (A) tail by deadenylases), cap structure removal (decapping or creating a 5'-mono-phosphorylated end that is recognized by Xrn), and processive mRNA turnover in the 5'-to-3' direction by an exoribonuclease (**Figure 4**; Mugridge et al., 2018; Nagarajan et al., 2013). In the cytoplasm of eukaryotic cells, 5'-to-3' decay is completed by Xrn1, while in the nucleus this process is completed by Xrn2 (Jones et al., 2012). However, in the cytoplasm of a plant cell, Xrn4 is used for 5'-to-3' RNA degradation (Kastenmayer & Green, 2000). All Xrns are extremely important for organisms, as lacking an Xrn will result in lethality or developmental defects (Nagarajan et al., 2013).

Cellular Xrns can also act to digest 5'-mono-phosphorylated viral transcripts in infected cells. However, Xrn digestion can be stalled by exoribonuclease resistant RNA structures, termed xrRNA, found in viral genomes. These xrRNAs stop Xrn progression and result in the production of viral degradation RNAs, herein termed small viral RNA (svRNA) (**Figure 8**; Steckelberg et al., 2018). These svRNAs can be non-coding RNAs and consist of viral 3'-UTRs or can be coding sequences for viral proteins (Gunawardene et al., 2019; Steckelberg et al., 2018).

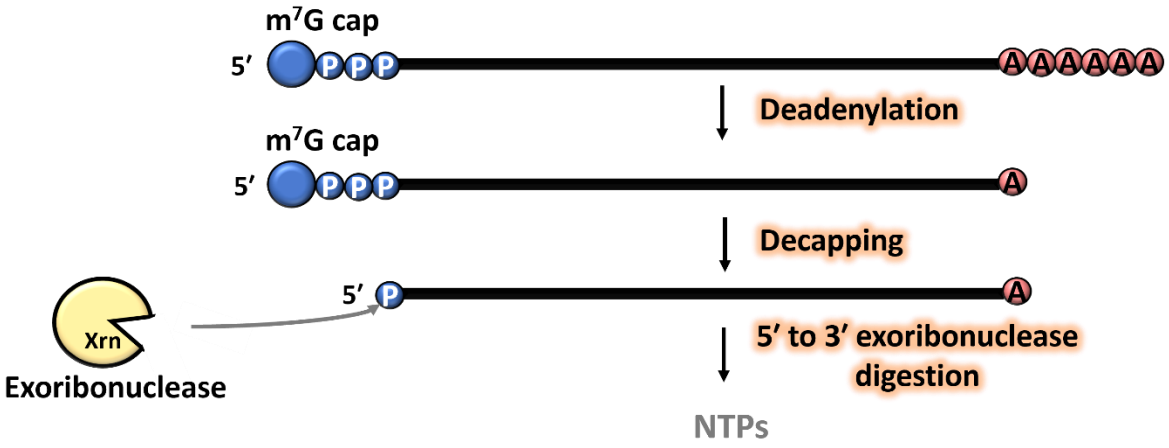


Figure 4. Overview of eukaryotic exoribonuclease (Xrn) degradation process. Cap structure is represented by large blue circle at the 5′-end of the mRNA labeled 7-methylguanosine cap (m^7G cap). The poly adenosine tail (poly (A) tail) is depicted by the string of red circles at the 3′-end of the mRNA labeled adenosine (A). Each step of the degradation process has labels highlighted in orange. First the mRNA is deadenylated by deadenylases, then it is decapped by decapping enzymes, leaving behind a 5′-monophosphate molecule that is recognized by Xrn (depicted by yellow circle labeled Xrn). Xrn facilitates 5′-to-3′ exoribonuclease digestion of the mRNA molecule into individual nucleoside triphosphates (NTPs) written in grey (Figure adapted from Mugridge et al., 2018).

1.5 Potato leafroll virus (PLRV)

Potato leafroll virus (PLRV) is a type species of the *Polerovirus* genus in the family *Solemoviridae* (formerly known as *Luteoviridae*). It is a virus that infects potato plants, but also has other agricultural hosts (Delfosse et al., 2021; Taliansky et al., 2003). PLRV is confined to the phloem of plants and replicates in the cytoplasm of the phloem and parenchyma cells (DeBlasio et al., 2018; Gill & Chong, 1975). Insects known as aphids, transmit the virus to other plants, in a circular, non-propagative manner, that is, the virus does not replicate in the aphid, but can be transmitted by the aphid (Dietzgen et al., 2016).

1.5.1 Gene expression

The PLRV genome is comprised of a positive-sense single-stranded RNA molecule that lacks a 5'-cap and 3'-poly(A) tail (**Figure 5**). Instead, at its 5'-end it carries a covalently attached viral protein, termed viral protein genome-linked or VPg and an RNA stem loop structure at its 3'-end (van der Wilk et al., 1997). The ORFs within the 5'-proximal half of the viral genome encode a suppressor of RNA silencing (P0) and an internally-located assistor of viral replication protein (Rap1) (**Figure 5**). As a suppressor of RNA silencing, P0 interacts with ubiquitin ligases that mediates the ubiquitination of ARGONATE1 (AGO1), which is a component of the RNA induced silencing complex (RISC) (Baumberger et al., 2007; Bortolamiol et al., 2007; Pazhouhandeh et al., 2006). Then, P0 and ubiquitinated AGO are loaded into endoplasmic reticulum (ER) associated vesicles that are sent to vacuoles for degradation (Michaeli et al., 2019). The Rap1 ORF is located within the P1 coding region and is translated from the genome using an IRES (Delfosse et al., 2021; Jaag et al., 2003). The function of Rap1 in poleroviruses is unknown, however it is proposed to be involved in regulating the transition from replication to transcription. Additionally, it is shown to interact with the P1 protein encoded by ORF1, located in the 5'-half of the PLRV genome (Jaag et al., 2003). The P1 ORF also encodes a polyprotein that includes a protease and the VPg. It is proposed that during VPg maturation the P1 protein is anchored to the intracellular membrane of the host, while the VPg domain binds to the viral genomic RNA, leading to proteolysis mediated VPg maturation (Delfosse et al., 2021; Prüfer et al., 1999). In some viruses, the VPg is involved in facilitating translation initiation (Duprat et al., 2002; Jiang & Laliberté, 2011), however it is not known if the PLRV VPg serves this role.

A -1 translational frameshifting event that involves a pseudoknot structure that causes the ribosome to slip back one nucleotide on a conserved heptanucleotide sequence is located in the overlapping region between ORF1 and ORF2 (**Figure 5**). This results in the ribosome shifting from the P1 ORF into the P2 ORF reading frame, generating the RdRP as a P1-P2 fusion (139 kDa) (**Figure 5**) (Delfosse et al., 2021).

ORFs in the 3'-proximal half of the genome encode MP, CP, and CP-RTD, with the latter being a C-terminal extension of the coat protein that is essential for aphid transmission (**Figure 5**). The CP-RTD product is produced in PLRV by translational readthrough involving a long-distance interaction between PRTE-DRTE (Chkuaseli & White, 2022; Xu et al., 2018). Three additional smaller proteins are also located in the 3'-region, P6 a protein of unknown function, P7, a protein potentially involved in viral regulation, and p3a (Delfosse et al., 2021). In PLRV p3a is a protein of only 5 kDa and is involved in systemic movement of the viral infection throughout the plant (Smirnova et al., 2015). It is translated using a non-AUG start codon AUA. Non-AUG start codons sometime have a downstream RNA stem loop (dSL) structure that helps with ribosome recognition, and examination of the region downstream of the AUA start codon by our laboratory suggested that a dSL may be present approximately 17 nts away from the non-AUG start codon (**Figure 6A**) (unpublished). The positioning of the dSL could assist the ribosome in stalling with its P-site over the AUA (**Figure 6B**; Kozak, 1990). One goal of this thesis was to explore if the putative dSL facilitated translation of p3a from the viral sg mRNA.

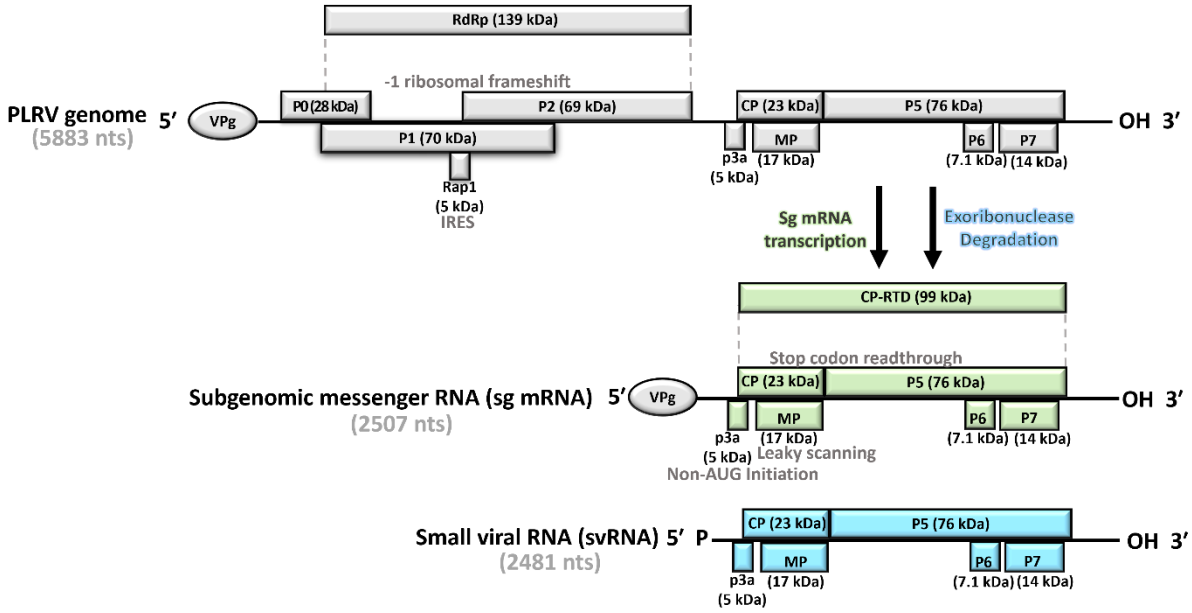


Figure 5. Outline of the PLRV genome, subgenomic messenger RNA (sg mRNA) and small viral RNA (svRNA). The open reading frames (ORFs) in the PLRV genome are shown by grey boxes. The ORFs in the sg mRNA are shown by green boxes and svRNA ORFs are shown by light blue boxes. Translational strategies used by the PLRV genome and sg mRNA are written in grey. The green highlighted words correspond to the production of sg mRNA and the blue highlighted words correspond to production of svRNA. Due to svRNA containing similar ORFs to sg mRNA, svRNA is predicted to produce CP-RTD and use the same translational strategies as sg mRNA.

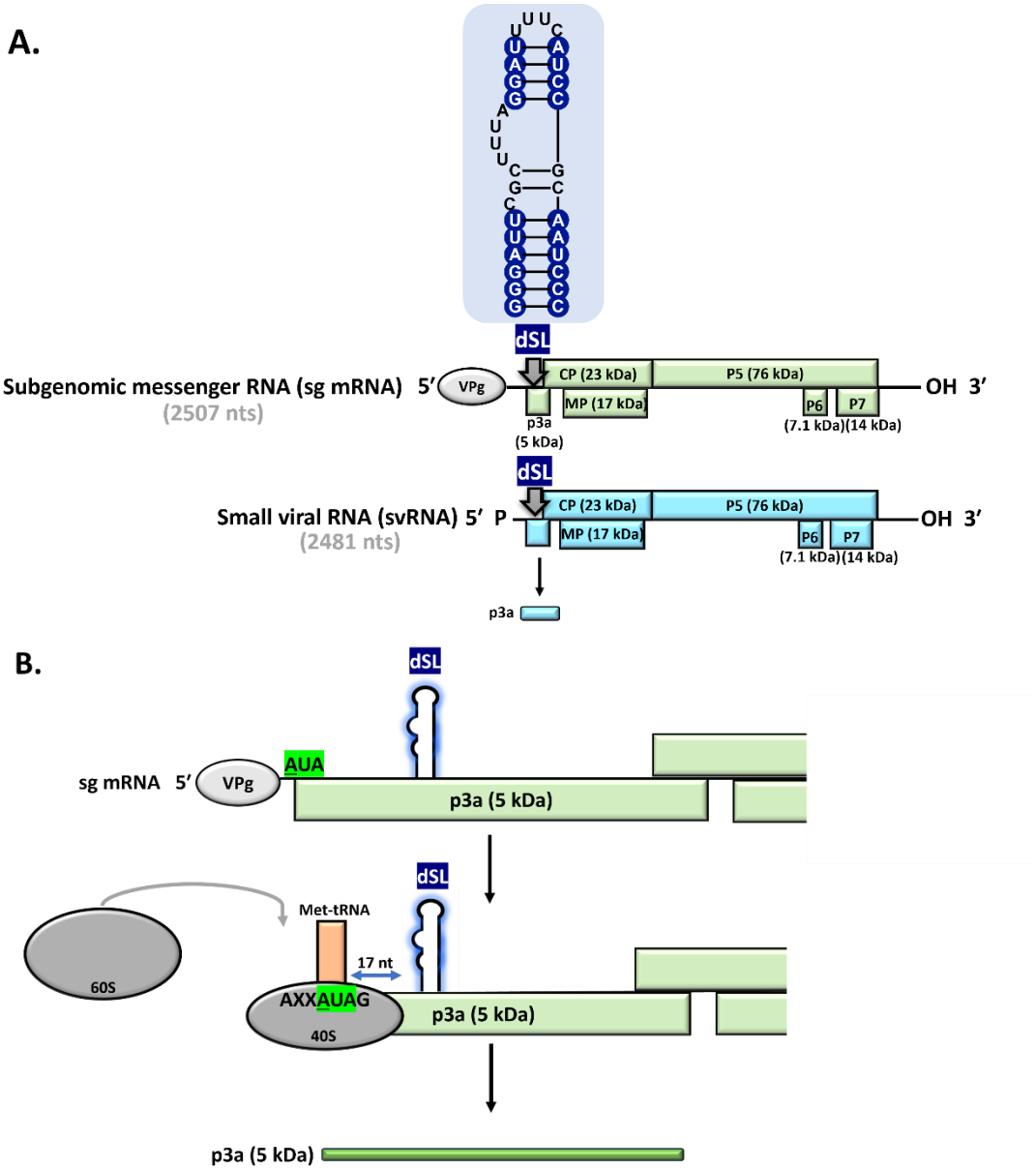


Figure 6. Downstream stem loop structure (dSL), location and potential mechanism. A. dSL structure highlighted in dark blue. Location of the dSL in the sg mRNA and svRNA are indicated by the thick grey arrows. The dark small light blue bar depicts the production of p3a. **B.** The potential mechanism of the dSL (highlighted in dark blue). The dSL promotes pausing of the ribosome (dark grey circles) at the non-AUG start codon (highlighted in green). This leads to entry of the Met-tRNA into the P-Site of the ribosome and 80S ribosome complex formation, which results in the production of p3a (dark green thin rectangle).

1.5.3 PLRV's subgenomic messenger RNA (sg mRNA)

Very little is known about PLRV's sg mRNA and how it is produced. The sequence at the 5'-terminus of the sg mRNA was characterized previously and found to be similar to the 5'-terminus of the genome (Miller & Mayo, 1991). However, what transcriptional mechanism **(Figure 3)** is used for its production remains unknown. Our laboratory identified a possible RNA structure in the PLRV genome that overlapped with the transcriptional start site of the sg mRNA **(Figure 7)** (unpublished). This structure could potentially be involved in mediating sg mRNA transcription and one of the goals of this thesis was to investigate this possibility.

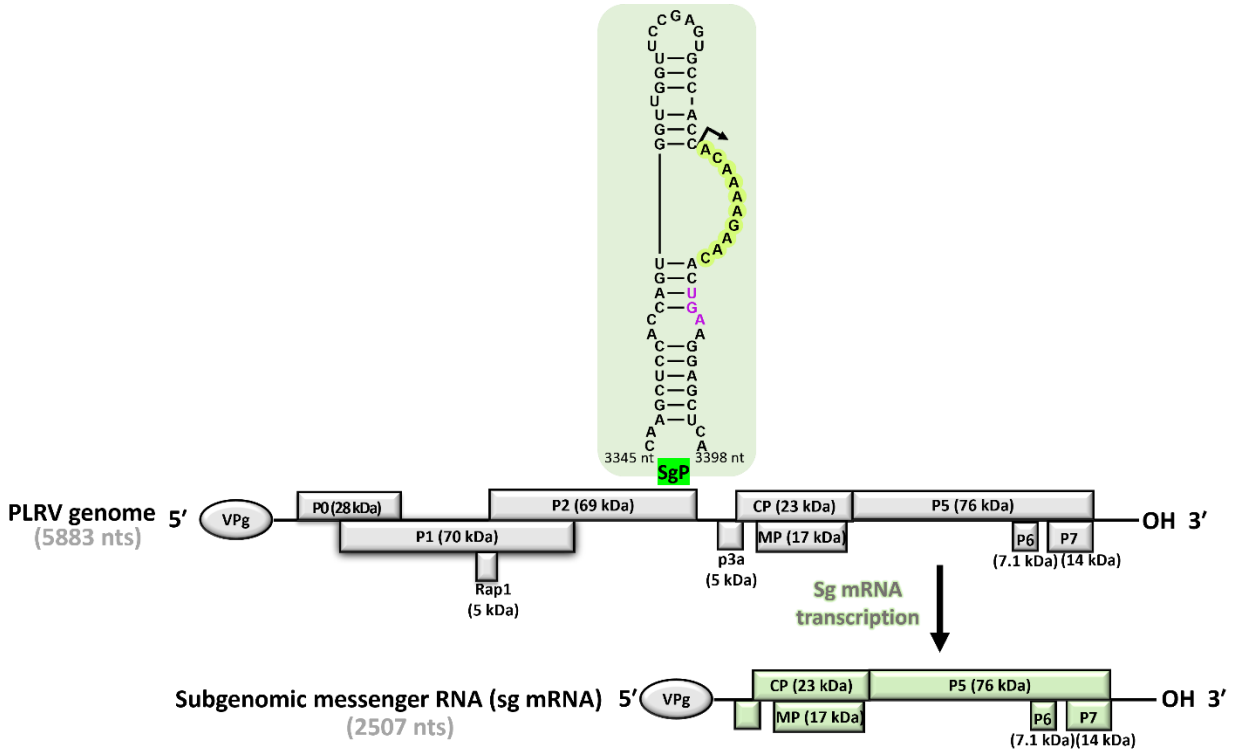


Figure 7. Subgenomic Promoter (SgP) element structure and location. The SgP is highlighted in green. It is potentially involved in the transcription of sg mRNA. The transcription start site is indicated by a small black arrow beside the green circled nucleotides, which is the promoter region. The purple nucleotides depict the stop codon for the RNA dependent RNA polymerase.

1.5.2 PLRV's small viral RNA (svRNA)

A viral RNA degradation product, termed small viral RNA (svRNA), was previously predicted to be generated from the PLRV genome (**Figure 5**) (Steckelberg et al., 2018). *In vitro* Xrn digestion assays showed that a degradation product was produced by an xrRNA structure that formed a pseudoknot and was located internally within the genome (**Figure 8**) (Steckelberg et al., 2018). The xrRNA structure was located 26 nts downstream of the sg mRNA transcription initiation site and the svRNA produced would be 26 nts shorter than the PLRV sg mRNA. Since the svRNA would also encode all of the viral ORFs in the sg mRNA, the possibility existed that it could also serve as a message for these viral proteins. However, it was not known if the svRNA could indeed serve as an mRNA or if it was produced during infections. Consequently, one of the goals of this thesis was to address these questions.

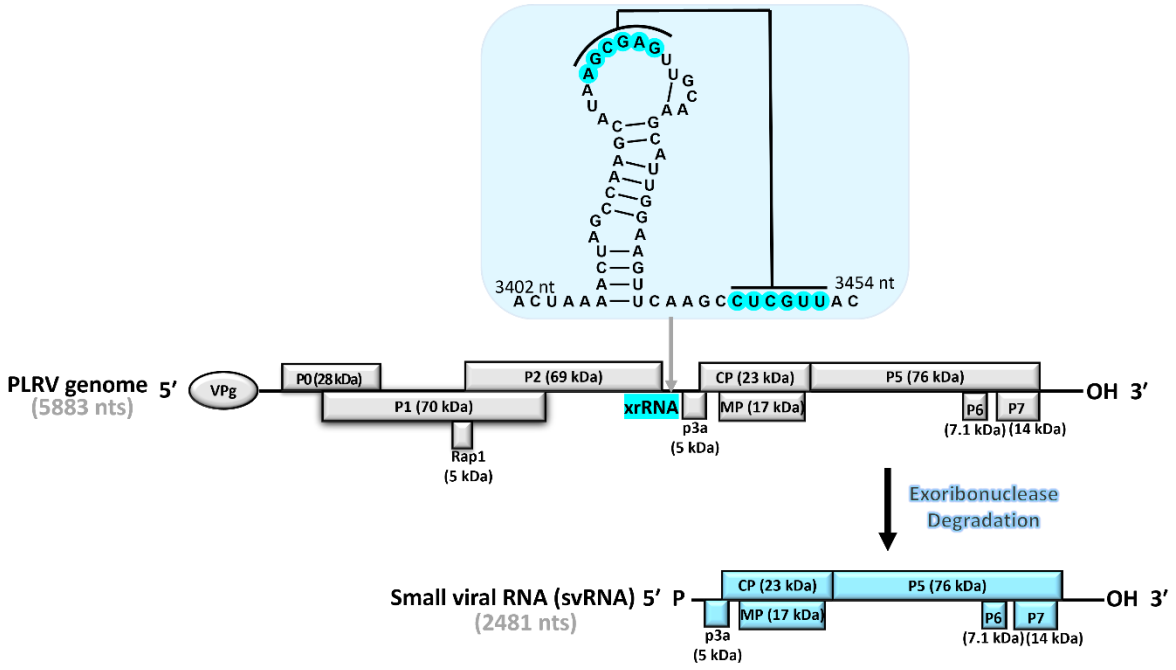


Figure 8. Exoribonuclease resistant RNA (xrRNA) structure location and proposed mechanism used to produce small viral RNA (svRNA). xrRNA structure is shown highlighted in light blue, location is indicated by grey arrow. The pseudoknot (pk) is depicted by black lines that show a base pairing interaction between the blue circled nucleotides. Its function is to produce svRNA via the exoribonuclease degradation process.

1.6 Research project objectives

The overall goal of this thesis was to investigate gene expression of PLRV, with a focus on how the proteins that are encoded 3'-proximally in the genome are produced. The project was divided into three parts, each with specific objectives. Part 1 involved investigating how the sg mRNA is transcribed. Part 2 looked at svRNA production and whether it was produced during PLRV infections. Part 3 examined translation of proteins from both sg mRNA and svRNA and the possible role of the dSL. The specific objectives for each of the three parts of this thesis are provided below.

PART 1: Investigating transcription of the sg mRNA

Objective 1: *In silico* analysis/comparative analysis of polerovirus sg mRNA promoter regions

Objective 2: Attempting to move the sg mRNA promoter out of the RdRp-coding region

Objective 3: Mutation of promoter sequences in its natural location

PART 2: Investigating svRNA production and accumulation

Objective 1: Demonstrating xrRNA-dependent generation of svRNA *in vitro* using yeast Xrn1

Objective 2: Determining if svRNA accumulates in pea protoplast infections

PART 3: Investigating protein translation from sg mRNA and svRNA

Objective 1: Characterization of sg mRNA translation products *in vitro*

Objective 2: Comparing sg mRNA and svRNA translation products *in vitro*

Objective 3: Determining if a downstream stem loop (dSL) enhances p3a production

2 Materials and Methods

2.1 Construct used for cloning

The construct used to clone all mutants was a PLRV genome cDNA (gene bank: KP090166.1) provided by Michelle Heck (Cornelle University). The provided construct included cDNA PLRV genome inserted into a pBin19 plasmid which was not suitable for in lab downstream research experiments such as protoplast infections. Therefore, the full cDNA PLRV genome was re-cloned into a pUc19 plasmid under the control of a T7 polymerase promoter to allow for *in vitro* transcription of the full-length PLRV RNA genome. The PLRV cDNA pUc19 construct (**shown in Figure 9**) was then used to clone other genomic mutants. This construct was also used to create PLRV sg mRNA and svRNA cDNAs of PLRV in pUc19 constructs (**Figure 9**), which were then used to design sg mRNA and svRNA mutants.

2.2 Oligonucleotide primers for polymerase chain reaction (PCR)

Primers were designed and ordered from Eurofins Genomics and used for PCR directed mutagenesis, sequencing, and northern blotting. All primers and northern blotting probes were diluted to a concentration of 10 µM before use. The PCR primers were designed to create PLRV genome, sg mRNA, and svRNA cDNA constructs, which were then used to create additional mutants. **Table 1** lists the primer sequences and descriptions involved in each PCR reaction. **Table 2** shows an overview of the overlapping PCR plan, which states the primers used, as well as the template, vector and restriction enzymes involved in generating each construct. The primers used for sequencing and northern blotting are found in **Table 3** (located in section 2.9) and **Table 4** (located in section 2.16), respectively.

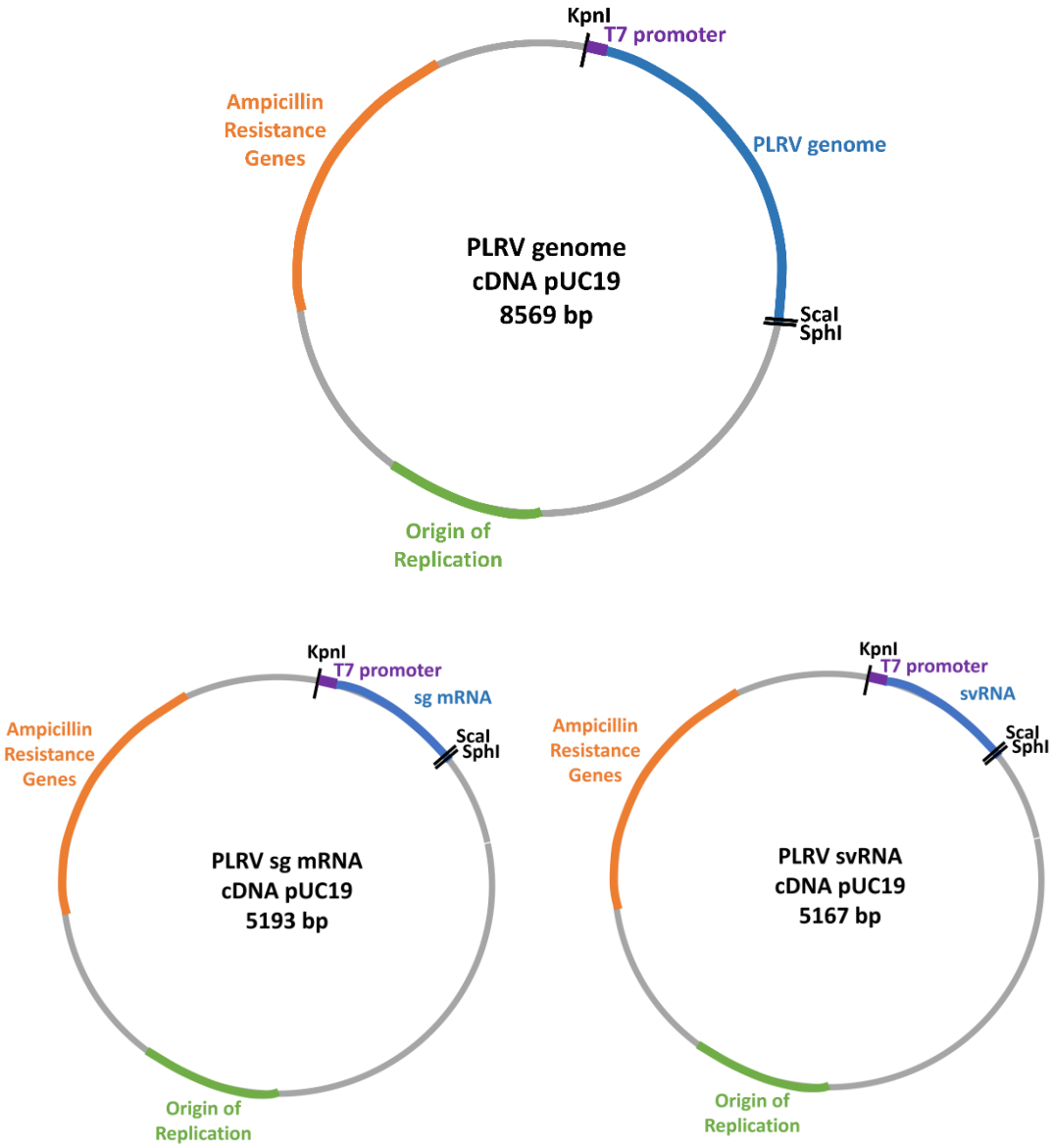


Figure 9. Simplified outline of wild type (wt), subgenomic mRNA (sg mRNA) and small viral RNA (svRNA) cDNA pUc19 constructs. Designed and created for downstream experiments. Restriction enzymes used for double restriction enzyme digestion reactions and linearization for cloning are depicted in black for each construct.

Table 1. Oligonucleotide primers used for polymerase chain reaction. The corresponding primer name and description is also provided. Red nucleotides represent addition mutations and yellow nucleotides represent substitution mutations. Black bold nucleotides indicate the T7 promoter required for *in vitro* transcription.

Primer Name	Oligonucleotide sequence (5' → 3')	Primer Description
pPL1	CCGGCGGTACC**TAATACGACTCACTATAG**ACAAAA GAATACCAGGAGAAATTGCA	Forward (55 nts)
pPL4	GGCGCGGCATGCCCTGCAGGAGTACTACACAACCC TGTAAGAGGATCCTGGCTACACA	Reverse (58 nts)
pPL5	CCGGCGGTACC**TAATACGACTCACTATAG**ACAAAA GAACACTGAAGGAGCTCAC	Forward (54 nts)
pPL6	CCGGCGGTACC**TAATACGACTCACTATAG**AAAACT AGCCAAGCATAAGCGAGTTG	Forward (55 nts)
pPL26	CGTTGCGTAGCTTCGAAATTTTGACACCATCCGC	Reverse (34 nts)
pPL27	CAACTGGTAGCCCGGGTTCTGTTCCAAAATCAG	Forward (33 nts)
pPL3Lpk1r	CTTCCAATGCTTGCAACTCGCTTATGCTTGGCTAG	Reverse (35 nts)
pPL3Lpk1fc	CGAGTTGCAAGCATTGGAAGTTCAAGCCT**GCG**TAC ATCAACCGG	Forward (44 nts)
pPL3Lpk2r	GCTTGAACTTCCAATGCTTGCAACT**GCA**TTATGCTT GGCTAG	Reverse (42 nts)
pPL3Lpk2f	GCAAGCATTGGAAGTTCAAGCCTCGTTACATCAACC GGAC	Forward (40 nts)
pPLpk3bpLcomprc	CTATTTTGTCCGGTTGATGTA**TGC**AGGCTTGAACTT CCAATGCTTGCAACT**GCG**TTATGCTTGGCTAG	Reverse (68 nts)
pPLpkcompaf	CATCAACCGGACAAAATAGATTATAAATTCTTAGC	Forward (35 nts)
pPL4pk1fc	CGAGTTGCAAGCATTGGAAGTTCAAGCC**AGCG**TAC ATCAACCGG	Forward (44 nts)
pPL4pk$_2$r	GCTTGAACTTCCAATGCTTGCAAC**AGCA**TTATGCTT GGCTAG	Reverse (42 nts)
pPLpk4bpcomprc	CTATTTTGTCCGGTTGATGTA**TGCT**GGCTTGAACTT CCAATGCTTGCAAC**AGCG**TTATGCTTGGCTAG	Reverse (68 nts)
pPLdSlmut1r	CCGGCTACTGAAAATGG**T**AT**A**GCGGATGAAAATCC TAAAGCGAA**A**CC**G**GCTAAGAATTTATAATC	Reverse (65 nts)
pPLdSlmut2r	CCGGCTACTGAAAATGGGATTGC**T**GA**A**GAAAA**A**CC TAAAGCGAATCCCGC	Reverse (50 nts)
pPLdSLmut3r	CCGGCTACTGAAAATGG**T**AT**A**GC**T**GA**A**GAAAA**A**CC TAAAGCGAA**A**CC**G**GCTAAGAATTTATAATC	Reverse (65 nts)
pPldSlKO1cr	CCGGCTACTGAAAATGGGATTGCGGATGAAAATCC TAAAGCGAA**AG**CCGCTAAGAATTTATAATC	Reverse (65 nts)
pPLdSLKO2cr	CAAAATATAAACCGG**CT**ACTGAAAATGGCTTTGCG GATGAAAATCCTAAAGC	Reverse (52 nts)

pPLdSLKO2f	CAGTAGCCGGTTTATATTTTGTTTACCTAAAGATTTCCTC	Forward (40 nts)
pPldSLcomr	CCGGCTACTGAAAATGGCTTTGCGGATGAAAATCCTAAAGCGAAAGCCGCTAAGAATTTATAATC	Reverse (65 nts)
pPLdSLf	CATTTTCAGTAGCCGGTTTATATTTTGTTTACCTAAAG	Forward (38 nts)
pPLsg1pf	GCAAAACCCCTCAGACACGATCGTGAGCTCGTTGCCAAGCTCCAC	Forward (45 nts)
pPLsg1pr	CTACCTATTTGCTAGTTTTAGTGAGCTCCTTCAGTGTTCTTTTG	Reverse (44 nts)
pPLsg1pcr2.1r	CGATCGTGTCTGAGGGGTTTTGCAAAGCCACCCTGATGGTGACTCTGAAGGATCC	Reverse (55 nts)
pPLsg1pcr2.2f	CTAAAACTAGCAAATAGGTAGACTCCGGATCAGAGCCTGGTCCAAGCCC	Forward (49 nts)
pPLsg2f	CGAAGCTACTCAGACACGATCGTGAGCTCGTTGCCAAGCTCCAC	Forward (44 nts)
pPLsg2r	GTTATCGTTGCGCTAGTTTTAGTGAGCTCCTTCAGTGTTCTTTTG	Reverse (45 nts)
pPLsg2pcr2.1r	CGATCGTGTCTGAGTAGCTTCGAAATTTTGACACCATCCGC	Reverse (41 nts)
pPLsg2pcr2.2f	CTAAAACTAGCGCAACGATAACACCTACCGCCAAGGTCACCC	Forward (42 nts)
pPLsg3f	GAACGGGACGCTCAGACACGATCGTGAGCTCGTTGCCAAGCTCCAC	Forward (46 nts)
pPLsg3r	GCTAATTGTAGCTAGTTTTAGTGAGCTCCTTCAGTGTTCTTTTG	Reverse (44 nts)
pPLsg3pcr2.1r	CGATCGTGTCTGAGCGTCCCGTTCAAGGAGTTGGCCCTCGCGAAAATG	Reverse (48 nts)
pPLsg3pcr2.2f	CTAAAACTAGCTACAATTAGCTTCCACGTTGAAGCGCCTACTG	Forward (43 nts)
pPLp3a/CPsvsgKOr	CTAAAGCGAATCCCGCTAAGAATTTATAATCTTTTTGTCCGGTTG	Reverse (46 nts)
pPLp3a/CPsvsgKOf	CTTAGCGGGATTCGCTTTAGGATTTTCATCCGCAATC	Forward (37 nts)
pPLsgsvATGr	CTAAAGCGAATCCCGCTAAGAATTTATAATCCATTTGTCCGGTTG	Reverse (45 nts)
pPLSY1r	GCTCCTTCAGTGTTCTTTTGTGGTAGCACTCGGAACCAACCACTGG	Reverse (46 nts)
pPLSY2r	GCTCCTTCAGTGTTCTTTTGTGGTGGCACTCGGAACTAACCACTGGTGGAGCTTGGC	Reverse (57 nts)
pPLSY3r	GAGCTCCTTCAGTGTTCTTTTGTGGTAGCACTCGGAACTAACCACTGGTGGAGCTTG	Reverse (57 nts)

pPLSY4r	GCTCCTTCAGTGTTCTTTTGTGGTGGCACTCGGAAC GAGCCACTGGTGGAGCTTGGC	Reverse (57 nts)
pPLSY5r	GCTCCTTCAGTGTTCTTTTGTCGTGGCACTCGGAAC GAGCCACTGGTGGAGCTTG	Reverse (55 nts)
pPLsgpf1	CAAAAGAACACTGAAGGAGCTCACTAAAACTAGCC AAG	Forward (38 nts)
pPLSB1r	GTGGCACTCGGAACCAACCACTGGTGCAGCTTGGC AACGAGCTCACGATCGTGTC	Reverse (55 nts)
pPLsgpf2	GTGGTTGGTTCCGAGTGCCACCACAAAAGAACACT GAAGGAGCTCAC	Forward (47 nts)
pPLsgpSB2f	CCGAGTGCCACCACAAAAGAACACTGAAGCAGCTC ACTAAAACTAGC	Forward (47 nts)
pPLSB3r	GTGGTGGCACTCGGAACCAACCACTGGTGCAGCTT GGCAACGAGC	Reverse (45 nts)
pPLsgpSB3f	GGTTGGTTCCGAGTGCCACCACAAAAGAACACTGA AGCAGCTCACTAAAAC	Forward (51 nts)
pPLSB4r	GTGGCACTCGGAACCAACCACTGGTGCAGTTTGGC AACGAGCTCACGATCG	Reverse (51 nts)
pPLSl2r	GTTCTTTTGTGGTGGCACTCGGAACCAACCACTGGT GGAGCTTGG	Reverse (45 nts)
pPLsgpSB5f	GAGTGCCACCACAAAAGAACACTGAAGCAGTTCAC TAAAACTAGC	Forward (45 nts)
pPLSB6r	GTTCTTTTGTGGTGGCACTCGGAACCAACCACTGGT GCAGTTTGGCAACGAGCTCAC	Reverse (57 nts)
pPLSGP1r	CTCGGAACCAACCACTGGTGGAGCTTGGCAACGAG C	Reverse (36 nts)
pPLSGP1f	CACCAGTGGTTGGTTCCGAGTGCCACCACCAAGGA GCACTGAAGGAGCTC	Forward (50 nts)
pPLSB7f	CGAGTGCCACCACAAAAGAACACTGAAGGAGCTTG CTAAAACTAGC	Forward (46 nts)
pPLSB8f	CGAGTGCCACCACAAAAGAACACTGAAGGTCGTCA CTAAAACTAG	Forward (45 nts)

Table 2. Overlapping PCR plan and cloning design for generation of all constructs and mutants. A. shows constructs designed using 3 step overlapping PCR. **B.** shows constructed designed using 4 step overlapping PCR.

A.

Constructs	Construct description/ purpose	Template (PCR1, PCR2)	PCR1	PCR2	Final PCR Template: PCR1 +PCR2	Vector	Enzymes Vector: wt
wt	creation of other cDNA puC19 constructs	wt°	pPL1 +pPL4			pUc19	KpnI + SphI
Sg		wt°	pPL5 + pPL4			pUc19	KpnI + SphI
Sv		wt°	pPL6 + pPL4			pUc19	KpnI + SphI
3Lpk1	xrRNA mutants	wt	pPL27 + pPL3Lpk1r	pPL3Lpk1fc + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
3Lpk2		wt	pPL27 + pPL3Lpk2r	pPL3Lpk2f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
3Lpkc		wt	pPL27 + pPLpk3bpLcomprc	pPLpkcompaf + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
4pk1		wt	pPL27 + pPL3Lpk1r	pPL4pk1fc + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
4pk2		wt	pPL27 + pPL4pk$_2$r	pPL3Lpk2f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
4pkc		wt	pPL27 + pPLpk4bpcomprc	pPLpkcompaf + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
sgm1	dSL mutants	sg	pPL5 + pPLdSlmut1r	pPLdSLf + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
sgm2		sg	pPL5 + pPLdSlmut2r	pPLdSLf + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
sgm3		sg	pPL5 + pPLdSLmut3r	pPLdSLf + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
sgKo1		sg	pPL5 + pPldSlKO1cr	pPLdSLf + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
sgKo2		sg	pPL5 + pPLdSLKO2cr	pPLdSLKO2f + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
sgC3		sg	pPL5 + pPldSLcomr	pPLdSLf + pPL26	pPL5 + pPL26	sg	Kpn1 + BstBI
svm1		sv	pPL6 + pPLdSlmut1r	pPLdSLf + pPL26	pPL6 + pPL26	sv	Kpn1 + BstBI

svm2		sv	pPL6 + pPLdSlmut 2r	pPLdSLf + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
svm3		sv	pPL6 + pPLdSLmut 3r	pPLdSLf + pPL26	pPL6 (F), pPL26 (R)	sv	Kpn1+ BstBI
svKo1		sv	pPL6 + pPldSlKO1c r	pPLdSLf + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
svKo2		sv	pPL6 + pPLdSLKO2 cr	pPLdSLKO2 f + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
svC3		sv	pPL6 + pPldSLcom r	pPLdSLf + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
sgp3aKO	confirmation of p3a identity mutants	sg	pPL5 + pPLp3a/CP svsgKOr	pPLp3a/CP svsgKOf + pPL26	pPL5 (f) pPL26 (r)	sg	Kpn1+ BstBI
sgp3aAUG		sg	pPL5 + pPLsgsvAT Gr	pPLp3a/CP svsgKOf + pPL26	pPL5 + pPL26	sg	Kpn1+ BstBI
svp3aKO		sv	pPL6 + pPLp3a/CP svsgKOr	pPLp3a/CP svsgKOf + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
svp3aAUG		sv	pPL6 + pPLsgsvAT Gr	pPLp3a/CP svsgKOf + pPL26	pPL6 + pPL26	sv	Kpn1+ BstBI
SB-1	sg promoter mutants	wt	pPL27 + pPLSB1r	pPLsgpf2 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-2		wt	pPL27 + pPLSl2r	pPLsgpSB2f + pPL26 (r)	pPL27 + pPL26	wt	SmaI + BstBI
SB-3		wt	pPL27 + pPLSB3r	pPLsgpSB3f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-4		wt	pPL27 + pPLSB4r	pPLsgpf2 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-5		wt	pPL27 + pPLSl2r	pPLsgpSB5f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-6		wt	pPL27 + pPLSB6r	pPLsgpSB5f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-7		wt	pPL27 + pPLSl2r	pPLSB7f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SB-8		wt	pPL27 + pPLSl2r	pPLSB8f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SGP-1		wt	pPL27 + pPLSGP1r	pPLSGP1f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SY-1		wt	pPL27 + pPLSY1r	pPLsgpf1 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SY-2		wt	pPL27 + pPLSY2r	pPLsgpf1 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SY-3		wt	pPL27 + pPLSY3r	pPLsgpf1 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI

SY-4		wt	pPL27 + pPLSY4r	pPLsgpf1 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
SY-5		wt	pPL27 + pPLSY5r	pPLsgpf1 + pPL26	pPL27 + pPL26	wt	SmaI + BstBI

°wt PLRV in pBin19 plasmid was used to create the PLRV wt inserted into pUc19 plasmid

B.

Construct	**Construct description / purpose**	**Template (PCR1, PCR2, PCR3)**	**PCR1**	**PCR2**	**PCR3**	**Final PCR (PCR4) Template: PCR1 +PCR2**	**Vector**	**Enzymes Vector: wt**
1sg	sg promoter insertion into CP-RTD ORF mutants	wt	pPLsg1pf + pPLsg1pr	pPL27 + pPLsg1pcr2.1r	pPLsg1pcr2.2f + pPL26	pPL27 + pPL26	wt	SmaI + BstBI
2sg		wt	pPLsg2f + pPLsg2r	pPL27 + pPLsg2pcr2.1r	pPLsg2pcr2.2f + pPL4	pPL27+ pPL4	wt	SphI + BstBI
3sg		wt	pPLsg3f + pPLsg3r	pPL27 + pPLsg3pcr2.1r	pPLsg3pcr2.2f + pPL4	pPL27 + pPL4	wt	SphI + BstBI

2.3 Overlapping polymerase chain reaction

The Q5 High-Fidelity DNA polymerase kit (New England Biolabs, NEB) and the GeneAMP PCR system 9700 (Applied Biosystems) was utilized for performing overlapping PCR. The components involved in the PCR reaction were 10 µl of 5x Q5 Reaction Buffer, 1 µl of 10 mM of deoxyribonucleotides triphosphate (dNTP), 2.5 µl of 10 µM of forward and reverse primers, 1 µg of template, 0.5 µl of Q5 High Fidelity DNA polymerase, 10 µl of 5x Q5 High GC Enhancer and diethyl pyrocarbonate-treated (DEPC) H_2O to reach a total volume of 50 µl.

The initial denaturation step of PCR reactions occurred at 98°C for 2 minutes to stimulate denaturation of the DNA template. This was followed by 30 cycles of denaturation at 98°C for 10 seconds to promote strand separation for primer annealing, which took place at 68°C for 30 seconds and then primer extension at 72°C. Duration of extension at 72°C was dependent on length of PCR fragment. After the 35 cycles were completed, a final extension was performed at 72°C for 5 minutes, which allows the Q5 High Fidelity DNA polymerase to fully complete synthesis of the PCR products. These same PCR conditions were used for all rounds of PCR.

2.3.1 Purification of PCR products through agarose gel extractions

The PCR products were purified via extraction of each sample from a 1% agarose gel. Each 50 µl PCR sample was mixed with 10 µl of 6x DNA loading dye [2 ml glycerol, 1 ml 1% bromophenol blue/1% xylene cyanol FF solution, 0.5 ml 0.5M EDTA (pH 8), 1.5 ml DEPC H_2O] to reach a total volume of 60 µl, that was loaded into the wells of the agarose gel together with DNA size markers. The gel was run at 150 V for 25 minutes in 1xTAE buffer to separate the PCR products. After making sure the PCR fragments had the expected size, the fragments were cut out of the gel. The cut-out gel pieces containing the PCR products were extracted from the gel using the GenepHlow Gel/PCR Kit from Geneaid. Following the kits instructions each cut-out gel piece was mixed with 500 µl of Gel/PCR Buffer and were incubated at 50°C for approximately 15 minutes to dissolve the gel. Once the gel was dissolved, each sample was placed in DFH columns and incubated at room temperature for 2 minutes to allow the DNA to bind to the columns filter membrane. The columns were then centrifuged at 10,621 x g for 30 seconds to remove the buffer solution. After centrifugation, the flow through was discarded, and the DNA was washed twice, once with 400 µl of W1 Buffer and a second time with 600 µl of wash buffer (with 100% ethanol added). Both washes were centrifuged at 10,621 x g for 30 seconds. To remove residual

wash buffer, the DFH column was centrifuged for an additional 3 minutes after discarding the flow through. Finally, the column was placed on a clean 1.5 ml microcentrifuge tube and 30 µl of DEPC H_2O heated to 60°C was added to the center of the column and incubated at room temperature for 2 minutes to allow the water molecules to bind to the DNA, before eluting by centrifuging at 15,294 x g for 3 minutes. These purified DNA samples were used as templates for the next or final round of PCR.

2.3.2 Purification of final round of PCR products

The final PCR products were purified with PCI (phenol chloroform isoamyl-alcohol; 25:24:1). The PCI mixture in each microcentrifuge tube was assembled with 50 µl of PCR reaction, 100 µl of PCI, 10 µl of 3M NaOAc (pH 5.2) and 40 µl of DEPC H_2O. The mixture was vortexed for 15 seconds then centrifuged at 21,130 x g for 5 minutes, at a temperature of 4°C. After centrifugation, the aqueous phase was transferred into another microcentrifuge tube containing 100 µl of PCI and was vortexed and centrifuged as before. Following the second centrifugation the aqueous phase was placed in a microcentrifuge tube containing 600 µl of 100% ethanol, for precipitation of the DNA at -20°C for 1-2 hours. Following the incubation, the samples were centrifuged at 21,130 x g for 15 minutes at a temperature of 4°C. Then the supernatant was discarded, and the DNA pellet was washed with 1 ml of 70% ethanol by centrifuging for 3 minutes. The supernatant was discarded again and briefly centrifuged to remove any residual ethanol. After removing all the 70% ethanol, the DNA pellet was dried at room temperature for 15 minutes, then dissolved in 30 µl of DEPC H_2O.

2.4 Restriction enzyme reactions

2.4.1 Double digestion reactions

Wt, sg, and sv PCR fragments and vectors were digested with restriction enzymes Kpn1, and SphI (NEB). The PCR digestion reactions were assembled with 5 µl of 10x rCutsmart Buffer (NEB), 5 µg of PCR fragment, 1 µl of KpnI (NEB, 20,000 U/ml), 1 µl of SphI (NEB, 20,000 U/ml) and DEPC H_2O to reach a total volume of 50 µl. The vector (wt) digestion reaction included the same reagents described previously with the addition of 1 µl calf-intestinal alkaline phosphatase (CIP) (NEB, QUICK CIP 5,000 U/ml). CIP was included in the digestion reaction only for the vector to dephosphorylate the 5'-ends of the vector to prevent self-ligation. Both digestion reactions of

the PCR fragments and the vector were incubated at 37°C overnight. Then they were purified using GenepHlow Gel/PCR Kit as described in section 2.3.1.

2.4.2 Sequential restriction enzyme digestion reactions

For the xrRNA mutants and subgenomic promoter (SgP) mutants, the same reagents and volumes were used as described in the above double digestion section with the exception of 2 µl of SmaI (NEB, 20,000 U/ml) added to the reaction first and incubated for 3 hours at 37°C. This was followed by the addition of 2 µl of BstBI (NEB, 20,000 U/ml) incubated for 3 hours at 65°C. Sequential digestion was also performed during the creation of the dSL mutants, using 2 µl of KpnI added to the reaction first for incubation at 37°C for 3 hours. This was followed by the addition of 2 µl of BstBI incubated for 3 hours at 65°C. Sg insertion in the CP ORF mutants involved the same sequential digestion conditions as described for the SgP mutants, except for mutants 2sg and 3sg using 2 µl of SphI and 2 µl of BstBI. Incubation duration and temperature are the same as described above. The vector reactions included the addition of 1 µl of CIP as described in the double digestion section. Also, the restriction enzymes used to digest each vector corresponded to the restriction enzymes used for the PCR fragments. This was done in order to create complementary digested ends for the next step, which was ligation. All restriction enzyme digested products were purified using GenepHlow Gel/PCR Kit as described in section 2.3.1.

2.5 Ligation

Once both the vector and PCR fragments were digested, using the restriction enzymes outlined in **Table 2**. The digested ends of the vector and PCR fragments were complementary to each other and bound together by a ligation reaction in order to form the constructs listed in **Table 2**. Assembly of the ligation reaction included 2 µl of digested and purified vector, 5 µl of digested and purified PCR fragment, 2 µl of 10x T4 DNA ligase buffer (NEB), 1 µl of T4 DNA ligase (400,000 U/ml) and DEPC H_2O to reach a total volume of 15 µl. A negative control was created for each vector, which was composed of digested vector with no added PCR fragment. The reaction was incubated at 16°C overnight.

2.6 Transformation of DNA plasmid into competent *E. coli* cells

For amplification and selection of successful ligation reaction the ligation products were transformed into competent *E. coli* DH5α cells on super Luria-Bertani broth (SLB) agar ampicillin (AMP) plates. The plasmids used (*i.e.* pUC19) contained an AMP resistance gene (**Figure 9**), that allowed for positive selection of the cells transformed with the ligated plasmids. Each ligation reaction was mixed with 50 µl of *E. coli* DH5α competent cells and incubated on ice for 15 minutes. To heat shock the cells the samples were placed in a 42°C water bath for exactly 90 seconds. The heat shock increased the pore size on the *E. coli* cell membrane, allowing the plasmids to travel inside the *E. coli* cells. Immediately after the heat shock the samples were placed on ice, which allowed the pore size of the *E. coli* cell membrane to decrease, trapping the plasmid inside the cell. Next 1 ml of SLB [Tryptone (32 g/L), Yeast Extract (20 g/L), NaCl (5 g/L), pH 7.5] without AMP was added to each sample and incubated at 37°C for 45 minutes to allow for cell recovery. After incubation, all microcentrifuge tubes containing the transformed *E. coli* cells were spun at 6,797 x g for 1 minute. Then most of the supernatant was discarded, leaving behind 50 µl. The pelleted cells for each sample were re-suspended by tapping and then spread on AMP SLB agar plates. The plates were incubated at 37°C for approximately 20 hours, to allow for selective growth of *E. coli* cells containing the DNA plasmids.

2.7 Amplification of DNA plasmid in super Luria-Bertani broth

The DNA plasmids were amplified by picking a single colony from the incubated AMP containing SLB agar plates and aseptically inoculating it, into 5 ml SLB containing 10 µl of AMP (50 mg/ml) for mini-preparations. For midi plasmid preparations 50 ml of SLB with 100 µl of AMP was used for each colony. The inoculated samples were incubated in an orbital incubator shaker (innova 4230 refrigerated incubator shaker), at 37°C, while shaking at 280 rpm for approximately 24 hours.

2.8 Extracting plasmid DNA from *E. coli* cells

2.8.1 Large-scale extraction of plasmid DNA (midi-preparation)

Midi plasmid preparations were completed for the large-scale extraction of the wt, sg mRNA and svRNA DNA plasmids. The bacterial cells were harvested by centrifuging in an RC5C Sorvall centrifuge at 4,303 x g for 5 minutes. After centrifugation, the supernatant was poured

off, leaving behind the pelleted bacterial cells. An alkali lysis method was used to lyse the bacterial cells. First the pellet was resuspended by vortexing in 150 µl of Solution I [Glucose (50 mM), Tris (25 mM; pH 8.0), EDTA (10 mM; pH 8.0)] then incubated at room temperature for 5 minutes. Next, 300 µl of Solution II was added to each sample. Solution II was made fresh with 1 ml of 0.2 N NaOH, 1 ml of 10% sodium dodecyl sulfate (SDS) and 8 ml of DEPC H_2O to make a total volume of 10 ml. After adding Solution II, the samples were mixed by inverting 20 times then incubated on ice for 5 minutes. Next, 250 µl of Solution III [60 ml potassium acetate (5 M), 11.5 ml glacial acetic acid, 28.5 ml DEPC H_2O] was added. The samples were mixed by inverting 20 times, then incubated at room temperature for 5 minutes.

The lysed bacterial cells were spun at 21,130 x g for 10 minutes to pellet them. The supernantant was removed and put into a new 1.5 ml microcentrifuge tube containing 600 µl of PCI. Then the samples were mixed through vigorous vortexing to denature remaining proteins. This was followed by a 15 minute centrifugation at 21,130 x g, at 4°C. After centrifugation, the supernatant was transferred to a new 1.5 ml microcentrifuge tube containing 600 µl of isopropanol. The samples were vortexed then incubated at -20°C for 15 minutes. After incubation, the precipitated DNA plasmids were pelleted by centrifugation at 21,130 x g for 10 minutes, at 4°C. Then the supernatant was discarded, and the pellet was left to dry at room temperature for 15 minutes to evaporate any residual isopropanol. The pellet was resuspended in 200 µl of DEPC H_2O. In order to digest any RNA present, the samples were treated with 4 µl of DNase-free RNaseA (10 mg/ml) by incubating at 37°C overnight. After the incubation, two rounds of PCI extraction were performed with 200 µl of PCI to extract the RNase from the sample. The PCI extraction involved vigorous vortexing for about 15 seconds, followed by 5 minute centrifugation at 21,130 x g, at a temperature of 4°C. After two rounds of PCI extraction the DNA plasmids were precipitated in 400 µl of 100% ethanol and incubated at -20°C for 15 minutes.

Following the precipitation, the DNA plasmids and smaller nucleic acids were pelleted by centrifuging for 10 minutes at 21,130 x g, at a temperature of 4°C. The supernatant was discarded, the pellet was dried at room temperature and resuspended in 100 µl of DEPC H_2O. To isolate the plasmid DNA, 300 µl of 20% PEG-8000 [polyethylene glycol; 20% PEG 8000, NaCl

(2.5M)] was added to the sample and pipette mixed, then placed on ice for 20 minutes to precipitate the plasmid DNA. Following the precipitation, the samples were centrifuged for 15 minutes at 4°C, the supernatant was removed, and 1 ml of 70% ethanol was added. The 70% ethanol was needed to wash the pellet and remove any salt or other impurities off the pellet. Finally, the samples were re-centrifuged at 21,130 x g for 5 minutes, the supernatant was discarded, and the pellet was dried at room temperature. The pellet was resuspended in 100-150 µl of DEPC H_2O. DNA plasmid concentration was determined by spectrophotometry $OD_{260/280}$ reading.

2.8.2 Small-scale preparations of plasmid DNA (mini-preparation)

Mini plasmid preparations were completed for small scale extraction of xrRNA, dSL, SgP insertion in the CP ORF and SgP mutants using Presto™ Mini Plasmid Kit (Geneaid). Bacterial cells were harvested by spinning down each overnight culture in the RC5C Sorvall centrifuge at 4,303 x g for 5 minutes and draining the supernantant. Alkali lysis method was used to lysis the bacterial cells. First, 200 µl of PD1 Buffer containing RNase A was added to the pellet to resuspend the bacterial cells and transfered into a 1.5 ml microcentrifuge tube. Next, 200 µl of PD2 Buffer was added to the mixture, mixed by inverting the samples 20 times and incubated at room temperature for 2 minutes. Then 300 µl of PD3 Buffer was added and mixed by inverting the samples 20 times.

The samples were centrifuged at 20,817 x g for 10 minutes. The supernatant was transferred into a PDH column, incubated at room temperature for 2 minutes to allow the DNA plasmids to bind to the column, then centrifuged at 10,621 x g for 30 seconds. The flow through was discarded and the PDH column was washed twice. Once with 400 µl of W1 Buffer and a second time with 600 µl of Wash Buffer (100% ethanol added). Both washes were centrifuged at 10,621 x g for 30 seconds. Then the column was spun for an additional 3 minutes at the same speed to remove any residual wash solution. Finally, the column was transferred onto a clean microcentrifuge tube, where 50 µl of DEPC H_2O was added to the center of the column to elute the DNA. After 2 minutes of incubation at room temperature, the column was centrifuged at 20,817 x g for 2 minutes. The purified plasmid DNA was quantified through spectrophotometry $OD_{260/280}$ readings.

2.9 Sequencing DNA clones

The extracted DNA plasmids sequences were verified through sequencing which was performed in The Centre for Applied Genomics at The Hospital for Sick Children. The sequencing results were verified for each construct before using them in downstream experiments. **Table 3** shows a list of all primers designed for sequencing.

Table 3. Primers used for sequencing reactions. In order to verify the sequences of each cDNA construct.

Primer Name	Oligonucleotide sequence (5' → 3')	Primer Description
pPL7	GGTACCTAATACGACTCACTATAGACAAAAGAATACCAGGAGAAATTGCA	Forward (50 nts)
pPL8	GAAAAGAGCGGCATATGCGG	Reverse (20 nts)
pPL9	GGACGCTTGCGAGGTTGATC	Forward (20 nts)
pPL10	GCCCATCACCAGTTCGTGTCTG	Forward (22 nts)
pPL11	GCAGGGTGTCTCAGAAAGCC	Forward (20nts)
pPL12	GGCATTCCCTATATCGCGTATG	Forward (22 nts)
pPL3	TCCGTCTTATGCCTGTCCGATGGCACTTTACTT	Forward (33 nts)
pPL13	GTGGTTGGTTCCGAGTGCC	Forward (19 nts)
pPL14	CGAGTCTATCAGACTGTCCGG	Forward (21 nts)
pPL15	GAGTATGCAAGCCGTCCCTATG	Forward (22 nts)
pPL16	GAGGGCCACATCTATATGG	Forward (19 nts)
pPL17	GGCAGTGGTTCTCTAACAGG	Forward (20nts)
pPL18	GCCTCGCTGACGAACTGAAG	Reverse (20nts)

2.10 Linearization and purification

Once the plasmid DNA sequences were verified, they were prepared for *in vitro* transcription by linearization using the ScaI restriction enzyme. The ScaI restriction site was located at the very 3'-end of all the viral cDNA sequences, thus generating an authentic 3'-end. The linearization reactions were assembled with 5 µl of 10x rCutSmart buffer, 20 µg of plasmid DNA, 1 µl of ScaI (20,000 U/ml, NEB) and DEPC H_2O to reach a final volume of 50 µl. The

linearization reactions were incubated at 37°C overnight. Then the linearized DNA plasmids were PCI purified as explained in Section 2.3.2. The samples were run in a 0.8% agarose gel alongside a DNA ladder and undigested cDNA construct to confirm successful linearization.

2.11 *In vitro* transcription

2.11.1 Capped *in vitro* transcription reactions

Capped *in vitro* transcription reactions were performed to obtain RNA for protoplast infections and *in vitro* translation assays. *In vitro* transcription reactions used a MessageMAX™ T7 ARCA-Capped Message Transcription Kit (CELLSCRIPT). The *in vitro* transcription reaction components include 2 µg of linearized DNA, 2 µl of 10X MessageMAX T7 Transcription Buffer, 8 µl of MessageMAX ARCA Cap/NTP PreMix, 2 µl of 100 mM DTT, 0.5 µl ScriptGuard RNase Inhibitor (40 U/µl), 2 µl MessageMAX T7 Enzyme Solution and DEPC H_2O to reach a total volume of 20 µl. The samples were incubated in a 37°C water bath for 1 hour. Then 1 µl of RNase free-DNase 1 (1 U/µl) was added to each sample, pipette mixed and incubated in the 37°C water bath for an additional 15 minutes to digest the DNA template. Following the incubation, 21 µl of 5 M NH_4OAc was added to each sample, pipette mixed and incubated on ice for 15-20 minutes to precipitate the RNA. The samples were then spun at 21,130 x g for 20 minutes and the supernatant was removed. To rid the RNA pellets of salts, 1 ml of 70% ethanol was added to the pellets and spun again at 21,130 x g for 15 minutes. After centrifugation, the 70% ethanol was removed, and the samples were briefly spun to remove residual ethanol. Then the pellets were air-dried at room temperature for 30 minutes and dissolved in 30 µl-40 µl of DEPC H_2O. The RNA transcript concentration was measured by spectrophotometry. RNA transcript quality and concentrations were verified by agarose gel electrophoresis. The RNA was stored at -20°C.

2.11.2 Uncapped *in vitro* transcription reactions

In vitro transcription reactions without a cap were performed to create PLRV genome, sg mRNA, and svRNA transcripts without a 5'-cap for use in *in vitro* translation reactions. These *in vitro* transcription reactions used a T7-FlashScribe™ Transcription Kit (CELLSCRIPT). The *in vitro* transcription reaction components included 2 µg of linearized DNA, 2 µl of 10X T7-FlashScribe transcription buffer, 1.8 µl of ATP (100 mM), 1.8 µl of CTP (100 mM), 1.8 µl of GTP (100 mM), 1.8 µl of UTP (100 mM), 2 µl of dithiothreitol (DTT, 100 mM), 0.5 µl of RNase Inhibitor (40 U/µl; ScriptGuard), 2 µl of T7-FlashScribe Enzyme Solution and DEPC H_2O to create a total volume of

20 µl. The samples were incubated in a 37°C water bath for 1 hour. Then 1 µl of RNase free-DNase 1 (1 U/µl) was added to each sample, pipette mixed and incubated in the 37°C water bath for an additional 15 minutes to digest the DNA template. RNA was precipitated using the same method ammonium acetate precipitation as described in the cap transcription section. It was also purified and dissolved in DEPC H_2O as described in the previous section.

2.12 Pea Protoplast Infections

2.12.1 Protoplast extraction and purification from mesophyll cells

To test the effects of the mutations on sg mRNA and svRNA accumulation *in vivo*, RNA was transfected into pea protoplasts. Protoplasts were isolated from pea plant (*Pisum sativum*) leaves that were picked from plants grown for 14 days. In order to extract protoplasts from the leaves, forceps were used to remove the lower epidermis layer. The peeled leaves were then placed peeled side down into a petri dish containing cellulase enzyme mixture [0.5 g Bovine serum albumin (BSA), 0.5 g Cellulase, 0.5 g macerozyme, 400 ml 10% mannitol]. The petri dish containing the peeled leaves were incubated in the dark at 26°C for 4 hours, while shaking at 60 rpm (innova 4230 refrigerated incubator shaker).

Following incubation, the contents in the petri dish were filtered with a miracloth (Millipore) into a 30 ml corex tube. The tube was then centrifuge in a Sorvall Legend T centrifuge at 80 x g for 5 minutes on brake setting 1, and most of the supernatant was removed. The pellet was resuspended, approximately 15 ml of 10% mannitol was added followed by the addition of approximately 15 ml of 40% sucrose at the bottom of the tube to create a density gradient. The tube was then spun at 80 x g for 5 minutes. The protoplast at the interphase of the sucrose and mannitol were transferred to a new corex tube and the same process of adding 10% mannitol, 40% sucrose and centrifuging was repeated.

2.12.2 Transfection of viral RNA transcripts into protoplasts

The protoplast interphase was transferred to a falcon tube and 10 µl of protoplasts were loaded into two chambers of a haemocytometer for quantification. Using a microscope, the average number of protoplasts were counted. Approximately, one million protoplasts were dispensed into small disposable glass tubes per infection. The tubes were centrifuged at 80 x g for 5 minutes and all supernatant was removed except for 100 µl. Most infections involved an inoculum containing 10 µg of RNA and DEPC H_2O to reach a volume of 15 µl was prepared in

microcentrifuge tubes. A negative control inoculum was also used, this involved 15 µl of DEPC H_2O and was called a mock infection. The protoplasts were resuspended by shaking, then the RNA inoculum was added to the protoplasts followed by shaking 10 times. Immediately after shaking the protoplasts and RNA were transfered into another glass tube containing 100 µl of PEG (40% PEG 1450/3 mM $CaCl_2$), followed by shaking 10 times. Then 800 µl of mannitol was added and tubes were shaken again 10 times. Finally, the glass tubes were incubated on ice for 15 minutes. After incubating the samples on ice, they were centrifuged at 80 x g for 5 minutes and about 900 µl of supernantant was removed. The pellets were resuspended in 800 µl of pea protoplast incubation medium [2% sucrose, gentamicin sulphate (1 µg/µl), 1x AOKI salts (2 mM KH_2PO_4, 10 mM KNO_3,10 mM $MgSO_4$, 100 mM $CaCl_2$, 1.6 mg/ml KI and 0.25 mg/ml $CuSo_4$-$5H_2O$), 12% mannitol] by shaking. The samples were prepared for incubation by covering the tops of the tubes with parafilm. All samples were incubated for 24 hours at 22°C under constant fluorescent light.

2.13 Total nucleic acid extraction from transfected protoplasts

About 800 µl of pea protoplast incubation medium was removed from each glass tube without disturbing the protoplasts settled at the bottom. Then 300 µl of RNA extraction buffer [40% DEPC H_2O, 40%:10X STE pH 8 (0.5 M Tris, 0.01 M NaEDTA, 1.0 M NaCl), 20%:10% SDS] and 300 µl of PCI was added to each glass tube. The protoplasts were resuspended by pipetting up and down 7 times to rupture the cell membranes. Then the contents were immediately transferred to microcentrifuge tubes to be vortexed for 30 seconds and spun at 21,130 x for 10 minutes at 4°C. Next, the aqueous layer was transferred to a new microcentrifuge tube containing 300 µl of PCI, vortexed for 30 seconds and spun again at 21,130 x g for 5 minutes at 4°C. The aqueous layer was transferred to another microcentrifuge tube containing 100 µl of 8 M NH_4OAc and 1 ml of 100% ethanol. All samples were vortexed and incubated at -20°C overnight to precipitate the nucleic acids. After precipitation, the samples were centrifuged at 21,130 x g for 30 minutes at 4°C to form the total nucleic acid pellet. The supernatant was discarded, 1 ml of 70% ethanol was added to the tubes and the tubes were centrifuged at 21,130 x g for 15 minutes to wash the pellet. The supernatant was discarded, and the pellets were left at room temperature to dry for 30 minutes. Dried pellets were dissolved in 20 µl of

DEPC H_2O for 5 minutes at room temperature. The samples were separated in a 2% agarose gel for 45 minutes at 150 V and levels of 25S ribosomal RNA bands were monitored to ensure even loading. The agarose gel was then used for northern blotting. The remaining nucleic acids samples were stored at -20°C until further use. Each set of mutants were tested in pea protoplast infections three-four times, independently.

2.14 Northern Blotting analysis of sg RNA and genomic RNA

2.14.1 Electrophoretic transfer

The total RNA from the agarose gel was electrophoreticaly transferred onto a positively charged nylon membrane (Hybond-N+). The material obtained for the transfer was a Hybond®-N+ cut to gel size, two pieces of Whatman filter paper cut slightly larger than the nylon membrane, two sponges and a transfer cassette. Prior to assembly the Whatman filter paper, nylon membrane and sponges were soaked in cold 0.5x TBE [Tris base, Boric acid, 0.5 M EDTA (pH 8.0)]. To facilitate the transfer the components were assembled in a transfer cassette, from bottom (black side) to top (white side): sponge, Whatman filter paper, RNA agarose gel (face down), nylon membrane, Whatman filter paper and sponge. The cassette was then placed in the electrophoretic transfer apparatus with the black side aligned with the black cathode. The apparatus was filled with 0.5x TBE and a magnetic stir bar was placed at the bottom. The transfer was run at 230 mA for approximately 5 hours, in 0.5x TBE, at a temperature of 4°C, with the magnetic stirrer on low to allow the circulation of the buffer. The next day the membrane was removed from the transfer apparatus, placed between two sheets of Whatman filter paper, and left on a vacuum sealed dryer for 2 hours at 80°C. After the membrane was dried the RNA was immobilized and ready for subsequent prehybridization and probing.

2.14.2 Prehybridization and hybridization of ^{32}P-labeled DNA oligo nucleotide probes

Prior to hybridization of the ^{32}P-labeled DNA oligonucleotide probes, each membrane was prehybridized with 10 ml of annealing mix [20x SSC, 1 M $NaPO_4$ (pH 7), 50x Denhardt's, 5 mg/ml yeast RNA, 10% SDS] to prevent non-specific binding. For prehybridization, the membrane was placed in a glass dish with a locked lid and left shaking at 55°C for 2 hours.

While the membrane was being prehybridized the DNA oligonucleotide probes were labeled at their 5'-end with γ^{32}P and purified through G-50 column chromatography. DNA oligonucleotides probes listed in **Table 4** were complementary to nucleotides at the 3'-end of

the PLRV genome. The 5'-end labelling reaction was assembled as follows: 2 µl of DNA oligonucleotide probe mixture was mixed with 3 µl of T4 polynucleotide (PNK) reaction buffer (10x, NEB), 5 µl of γ^{32}P ATP (activity of 1 µCi, PerkinElmer), 0.5 µl of T4 PNK (10,000 U/ml, NEB) and DEPC H_2O to reach a total volume of 10 µl per reaction. The oligonucleotide probe mixture was placed at 37°C for 45 minutes. During this time, the γ^{32}P was transferred from ATP to the 5'-OH termini of the oligonucleotides. Following the labeling process, the probes were purified by column chromatography, using a Sephadex G-50 spin column. The column was prepared with 800 µl of Sephadex® G-50 added to an EZ-10 spin column (Biobasic Inc) and spun at 718 x g for 3 minutes. The probe mixture was added to the assembled G-50 spin column along with DEPC H_2O, to form a total volume of 50 µl and the probe mixture was spun at 718 x g for 2 minutes, to allow the probes to filter through the column into a new microcentrifuge tube. The DNA oligonucleotide probes were complementary to the 3'-end of the genomic RNA, sg mRNA, and svRNA. Probes were added to the membrane with shaking overnight at room temperature.

2.14.3 Washing the nylon membrane and radioactivity quantification

Following hybridization of the probes to the nylon membrane, a series of low and high stringency washes were completed using 2x SSC/0.1% SDS [890 ml H_2O, 10 ml 10% SDS, 100 ml 20x SSC to make a total volume of 1000 ml]. Two low stringency washes were performed with 50 ml of 2x SSC/0.1% SDS added to a glass dish containing the nylon membrane and shaken at room temperature for 20 minutes. Then two high stringency washes were carried out with 50 ml of 2x SSC/0.1% SDS that was warmed to 50°C added to the glass dish and shaken at 50°C for 30 minutes. After completing all the washes, the membrane was air dried, covered in plastic wrap and placed in a cassette. The membrane was then exposed for 1 week to a phosphor screen in the dark. Following exposure, the screen was analysed in a Typhoon FLA 9500 variable mode imager (GE Healthcare). The intensity of each detected RNA band was quantified using Quantity One® software (Bio-Rad). The ratio was found between sgRNAs and genomic RNA raw values. Wt sgRNAs/genomic RNA ratios were made equal to 100%. SgRNAs/genomic RNA ratios of all the mutants were calculated relative to wt. The infections were carried out three-four times independently and the average sgRNAs/genomic RNA ratios and standard error were calculated.

Table 4. Northern blotting probes sequences. Reverse primers used as probes that are complementary to RNA for Northern blotting.

Primer Name	Oligonucleotide sequence (5' → 3')	Primer Description
pPL23	CAGCTATGTTTCTATCTACTTGGGGAGGTGGC	Reverse (32 nts)
pPL24	CCTTTCGGCTTTCGTTCCGCTTAATGTCCGGC	Reverse (32 nts)
pPL25	GGATCCTGGCTACACAGTCGCGTCTTTCGACGGTC	Reverse (35 nts)
pPL28	GTTGTCTTTCCTGCGTTTGTATCGGGGTTTCGTCCCTTG	Reverse (39 nts)
pPL29	CTGTTTACCGAACCAGCATCGGATACGTCGTCAG	Reverse (34 nts)
pPL30	TCGGAGAAAAGTCAGTAACAGTTCTTGAATTGCCGG	Reverse (36 nts)
pPL31	CATTTTCCTCTCTCTTCTCCAGCATACTTC	Reverse (30 nts)

2.15 Xrn1 degradation assay

2.15.1 *In vitro* transcription reaction introducing a 5'-mono-phosphate

In vitro transcription reactions that generated 5'-mono-phosphorylated transcripts were used for yeast Xrn1 degradation assays. All reagents and volumes were the same as described in section 2.11.2, except 0.6 µl of GTP (100 mM) and 1.2 µl GMP (0.5 M) were added to each reaction. Incubation duration, temperature and RNA precipitation and purification is the same as described previously for other *in vitro* transcription reactions.

2.15.2 Yeast Xrn1 *in vitro* degradation assay

In vitro transcribed 5'-mono-phosphorylated transcripts were required for yeast Xrn1 degradation assays because yeast exoribonuclease (Xrn1) 5'-to-3' digestion requires a 5'-mono-phosphate. Reactions containing yeast Xrn1 consisted of 1 µl of 10x buffer 3 (NEB), 0.5 µl RNase inhibitor (40 U/µl, Biobasic Inc), 1 µl of Xrn1 (NEB, diluted to a concentration of 0.25 U/µl), 2 µg of 5'-mono-phosphorylated RNA, and DEPC H_2O to reach a total volume of 10 µl. Negative controls involved reactions with no Xrn1. All reactions were incubated at 37°C for 30 minutes to allow 5'-to-3' digestion of the RNA by Xrn1. After digestion, the 10 µl reactions were prepared for RNA precipitation by mixing with 1 µl of glycogen (20 mg/ml), 15 µl 3 M NaOAc (pH 5.2) and DEPC H_2O to reach a volume of 150 µl. Then 150 µl of PCI was added, this was followed by vortexing and centrifuging at 21,130 x g for 10 minutes at 4°C. Finally, the RNA was precipitated

in 400 µl of 100% ethanol by incubating at -20°C for 1 hour. After precipitation, the RNA was washed and dried as described in previous sections and dissolved in 20 µl of DEPC H_2O. The digested RNA fragments were separated from undigested fragments by gel electrophoresis in a 1.4% agarose gel stained with ethidium bromide (EtBr), run at 150 V for 20 minutes. A marker svRNA transcript was used to verify the identity of the full-length svRNA.

2.16 *In vitro* translation in a wheat germ extract (wge) system

Uncapped and capped RNA transcripts were used for *in vitro* translation assays. Translated proteins were radioactively labeled with S^{35}-Met, separated by SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel) and visualized after exposure to a phosphorimager. *In vitro* translation reactions were completed using an *in vitro* translation wheat germ extract (wge) kit from Promega. Reactions included 0.6 µl of RNase inhibitor (40 U/µl, Biobasic Inc), 10 µl wge, 1.6 µl of amino acid mix (-Met), 1.6 µl of 1 M KOAc, 0.5 µl of S^{35}-Met (activity 5 µCi, PerkinElmer), 0.5 pmols of RNA transcript and DEPC H_2O to reach a total volume of 20 µl. The reactions were incubated at 25°C for 1 hour. The proteins were then denatured by mixing 10 µl of the reaction with 10 µl of 2x SDS-PAGE loading dye [1.25 ml of 1 M Tris-HCl (pH 6.8), 2 ml of 100% glycerol, 4 ml of 10% SDS, 0.2 ml of 1% bromophenol blue, 1 ml of 100% BME and DEPC H_2O added to a volume of 10 ml] and boiling at 90-99°C for 5 minutes. A volume of 10 µl was loaded into the SDS-PAGE (12% resolving, 5% stacking gel) and run in 10x SDS-PAGE running buffer [10 g of SDS, 30.3 g of Tris, 144.1 g of glycine and 700 ml of ddH_2O to reach a total volume of 1000 ml]. Gels were run at 200 V for 30 minutes, which allowed for detection of the 5 kDa p3a in the gel.

Once the gels finished running, they were dried at 80°C for 1 hour, followed by exposure onto a phosphorimager screen for 1 week. Screens were then scanned using the Typhoon FLA 9500 variable mode imager (GE Healthcare) and protein accumulation was quantified using Quantity One® software (Bio-Rad). Accumulation levels of mutants were determined as a ratio of the level of their wt counterparts. Wt levels were made to equal 100% and all proteins translated in the mutant reactions were calculated relative to that of the wt. The *in vitro* translation reactions for each set of mutants were carried out three times, independently, the average protein ratios and standard error were calculated.

3 Results

3.1 Investigating transcription of the sg mRNA

3.1.1 Objective 1: *In silico* comparative analysis of polerovirus sg mRNA promoter regions

As stated previously in the introduction, nothing is currently known about sequences or structures that control transcription in poleroviruses. The first step in identifying PLRV's sg mRNA promoter (SgP) involved examining multiple polerovirus sequences surrounding the sg mRNA initiation site. This comparative structural analysis suggested the presence of conserved RNA secondary structures; formed through base pairing of the blue and yellow highlighted sequences (**Figure 10A**). It was noted that the yellow pairing was more conserved than the blue pairing. Comparative analysis of polerovirus sequences also revealed a conserved linear sequence corresponding to the 5′-end of the sg mRNA (green). Thus, based on conservation, the SgP was proposed to involve both secondary structures (blue and yellow) and a linear sequence (green) (**Figure 10A**). Note that these sequences and structures overlap with the C-terminal coding region of the P1/P2 RdRp (the RdRp stop codons are in magenta) (**Figure 10A**).

RNAStructure folding software (Reuter et al., 2010) was used to generate plus- and minus-strand sg mRNA folds of the promoter regions and examples are shown in **Figure 10B**. The comparisons suggested that similar structures can potentially form in both the plus- and minus-strands. Therefore, it is possible that, if the interactions (yellow and blue) are important for sg mRNA transcription, they may occur in the plus- and/or minus-strands. Overall, the conservation observed suggests that the sequences and structures are likely important transcriptional regulatory elements, and thus were investigated.

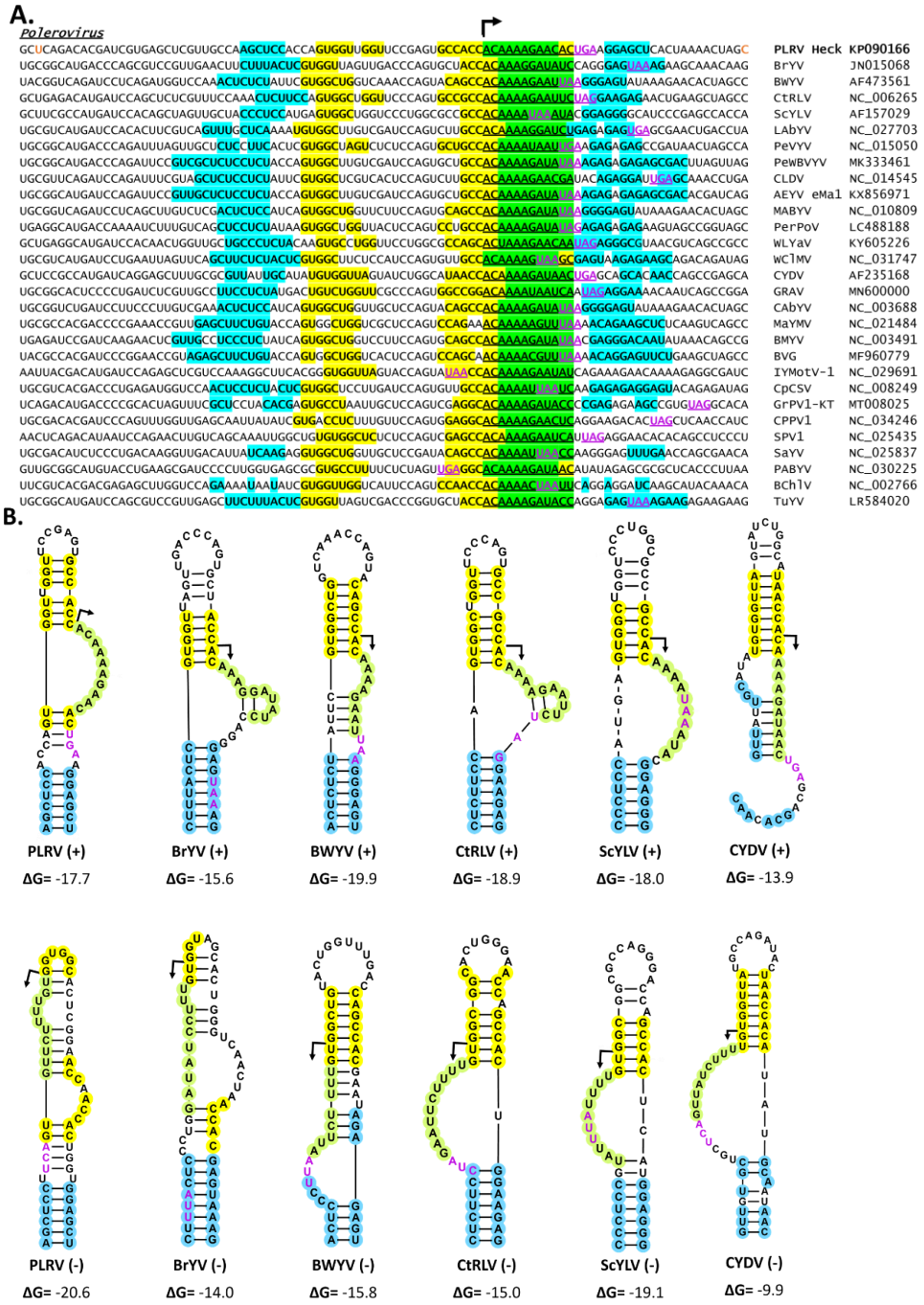


Figure 10. Comparative analysis of polerovirus subgenomic RNA promoter (SgP) regions. A. Pairs of blue and yellow highlighted viral sequences are predicted to base pair, the green highlighted and underlined sequence represents the promoter region. The black arrow at the top above PLRV indicates the transcription start site. Nucleotides in magenta represents the RdRp ORF stop codons. Single orange nucleotides found at the 5' and 3'-ends of the PLRV sequence indicate the region used to create SgP insertion mutants in figure 11. Accession numbers are provided on the right. **B.** *RNAStructure* local folds of equivalent SgP element structures [plus (+) and minus (-)] shown for Brassica yellows virus (BrYV, JN015068), Beet western yellows virus (BWYV, AF473561.1), Carrot red leaf virus (CtRLV, NC_006265), Sugarcane yellow leaf virus (ScYLV, AF157029.1), Cereal yellow dwarf virus (CYDV, AF235168). Colour scheme follows that in panel A. Black arrow indicates the transcription start site and corresponding ΔG for each local fold are given below each structure.

3.1.2 Objective 2: Attempting to move the sg mRNA promoter out of the RdRp-coding region

Sequence analysis indicated that the sg mRNA promoter sequences are located within the RdRp ORFs, and this would limit the introduction of promoter nucleotide substitutions to only wobble positions. An alternative method of testing the promoter region involves moving a copy of the promoter region to a different non-essential region of the viral genome, such as the CP-RTD ORF (Wang & Simon, 1997; Wong et al., 2017; Wu et al., 2010). In the new location, the sequence could be mutated without worrying about modifying important overlapping coding sequence. To test this approach with PLRV, a segment containing the potential SgP was moved to three different positions in the CP-RTD ORF (**Figure 11A**), creating genomic mutants 1sg, 2sg, and 3sg. If active, placement of the SgPs in three different positions in the CP-RTD ORF would produce additional sg mRNAs of various lengths, as indicated in **Figure 11A**. Note that CP-RTD is not required for protoplast infections.

PLRV genomes containing individual SgP insertions were tested by transfection into pea protoplasts. Following incubation, total nucleic acids were prepared and analyzed by northern blotting. All mutant genomes were viable and accumulated in infections. Mutants 1sg, 2sg, and 3sg showed similar migration patterns compared to wt, however no additional sg mRNA of the sizes predicted were observed (**Figure 11B**). The wt sgRNAs band (**labeled as wt sg mRNA (2507 nt) and svRNA (2481 nt) in Figure 11B**) in these mutants migrated a bit slower than the sgRNAs in the wt PLRV infection. This slower mobility was caused by the extra SgP sequences added in the mutants, that increased the size of the wt sg mRNA. Since no additional sg mRNA bands were observed in the infections with the SgP insertion mutants, the attempt to move an active SgP to a non-essential coding region was not successful.

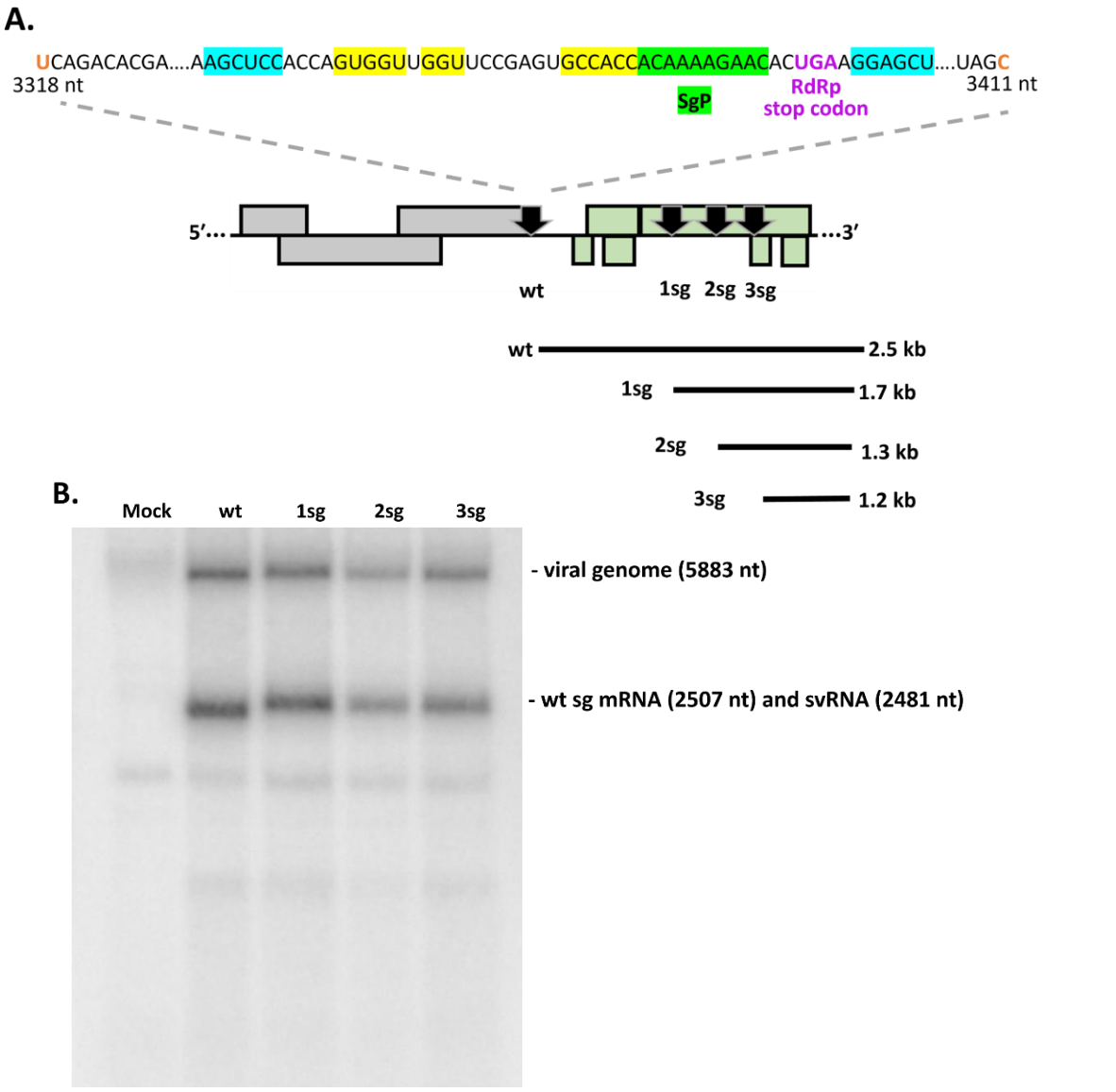


Figure 11. Subgenomic (sg) insertion mutants diagram and northern blot. A. Diagram of PLRV genome sg insertion mutations. The inserted putative SgP sequence is shown above the PLRV genome. The blue and yellow substructures, as well as sg promoter (SgP) and RdRp stop codon are present within the inserted sequence. Black arrows pointing downward indicate the three different locations of the insertions. The predicted sg mRNA that would be generated from each mutant is shown below the genome, with accompanied fragment length. **B.** Northern blot SgP insertion mutants' infections. A representative northern blot shows viral RNA accumulation from pea protoplast infections. Samples were separated in a 2% agarose gel at 150 V for 45 minutes, then transferred onto a positive charged nylon membrane. Viral RNAs were detected with 3'-terminal oligonucleotide probes labeled with $\gamma^{32}P$.

3.1.3 Objective 3: Mutation of promoter sequences in its natural location

Since the transplanted promoter insertion mutants were not successful, analysis of the SgP in PLRV was carried out in its natural RdRp-coding region. To accomplish this, compensatory and disruptive mutations at wobble positions were created in the yellow (SY mutants) and blue (SB mutants) interacting nucleotides, as well as substitutions in the green linear region (SGP-1 mutant).

3.1.3.1 SY mutants (yellow interaction)

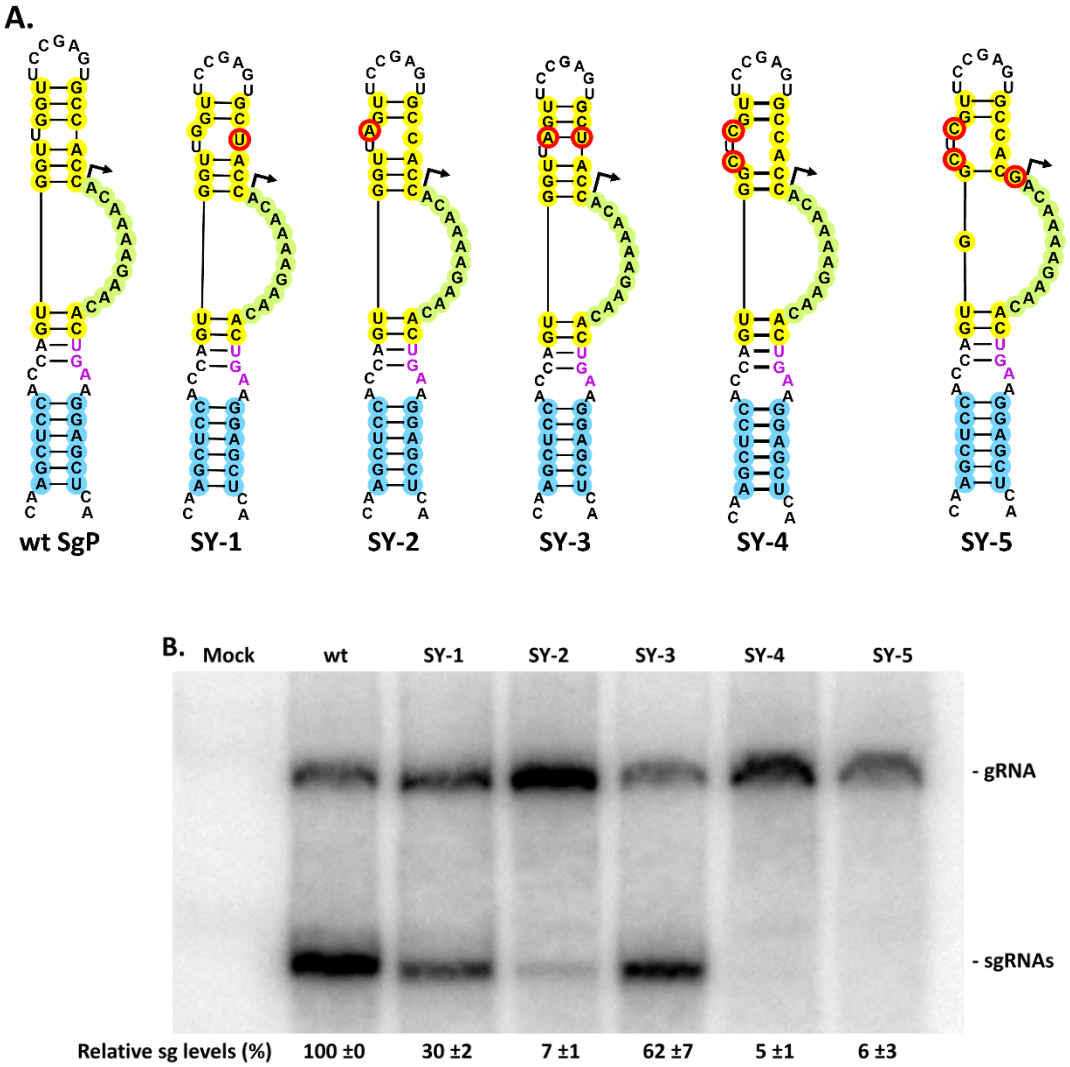
The yellow interaction was investigated first with a set of compensatory mutations (SY-1, SY-2 and SY-3) and two additional mutants predicted to efficiently disrupt the interaction (SY-4 and SY-5). Disruptive compensatory mutants SY-1 and SY-2 were predicted to destabilize the yellow structure and reduce accumulation of sgRNAs (*i.e.* comigrating sg mRNA and svRNA band in northern blot, **Figure 12B and D**), while restorative compensatory mutant SY-3 was predicted to regenerate the yellow structure and restore wt accumulation of sgRNAs (**Figure 12A**). SY-4 and SY-5 were designed to destabilize the yellow structure with several substitutions (**Figure 12A**).

All SY mutants were tested through transfection into pea protoplasts and detection of viral RNA through northern blotting (**Figure 12B**). The results showed a decrease in relative sgRNAs accumulation for SY-1 (30%) and an even greater decrease in sgRNAs accumulation for SY-2 (7%). The compensatory SY-3 was able to partially regain accumulation levels of sgRNAs to 62%, consistent with formation and functional importance of the yellow interaction. The two disruptive mutants SY-4 and SY-5 showed the lowest accumulation of sgRNAs, at 5% and 6%, respectively. Analysis of the yellow base pairing interaction indicates that its formation is important for sg mRNA transcription in PLRV.

3.1.3.2 SB (blue interaction) and SGP-1 (green sequence) mutants

Mutants tested for the blue interaction consisted of two sets of compensatory mutants (SB-1, SB-2, SB-3 and SB-4, SB-5, SB-6, respectively) and mutants that either further stabilized (SB-7) or disrupted (SB-8) the interaction (**Figure 12C**). Additionally, a mutant (SGP-1) was also created in which the linear green sequence was modified with three substitutions (**Figure 12C**).

All mutants were tested by transfection into protoplasts and northern blot analysis. For the blue interaction, both sets of compensatory mutants, either disruptive or restorative, showed a range of accumulation levels, 42% to 79%, which does not correlate with the interaction being important (**Figure 12D**). Therefore, the results with the compensatory mutants did not support a role for the blue interaction in sg mRNA transcription. Similar results were also observed with the individual SB-7 (61%) and SB-8 (42%) mutants that were predicted to further stabilize or to destabilize the blue interaction, respectively. However, compensatory set disruption mutants visually show a decrease in genomic RNA accumulation in SB-1, SB-2 and SB-4, SB-5 (**Figure 12D**). While compensatory mutants that restore the blue interaction SB-3 and SB-6 regain genomic RNA accumulation. Similar results were seen in the additional mutant's suggesting that the blue interaction may be important for genomic RNA accumulation. The results for SGP-1, that modified the linear green sequence, showed a decrease in relative sgRNAs accumulation (29%) (**Figure 12D**). In summary, the results suggest that the blue interaction is not required for sg mRNA transcription but may play a role in genome accumulation and that the green sequence is important for transcription.


A.
wt SgP
SY-1
SY-2
SY-3
SY-4
SY-5
B.
Mock
wt
SY-1
SY-2
SY-3
SY-4
SY-5
- gRNA
- sgRNAs
Relative sg levels (%)
100 ±0
30 ±2
7 ±1
62 ±7
5 ±1
6 ±3

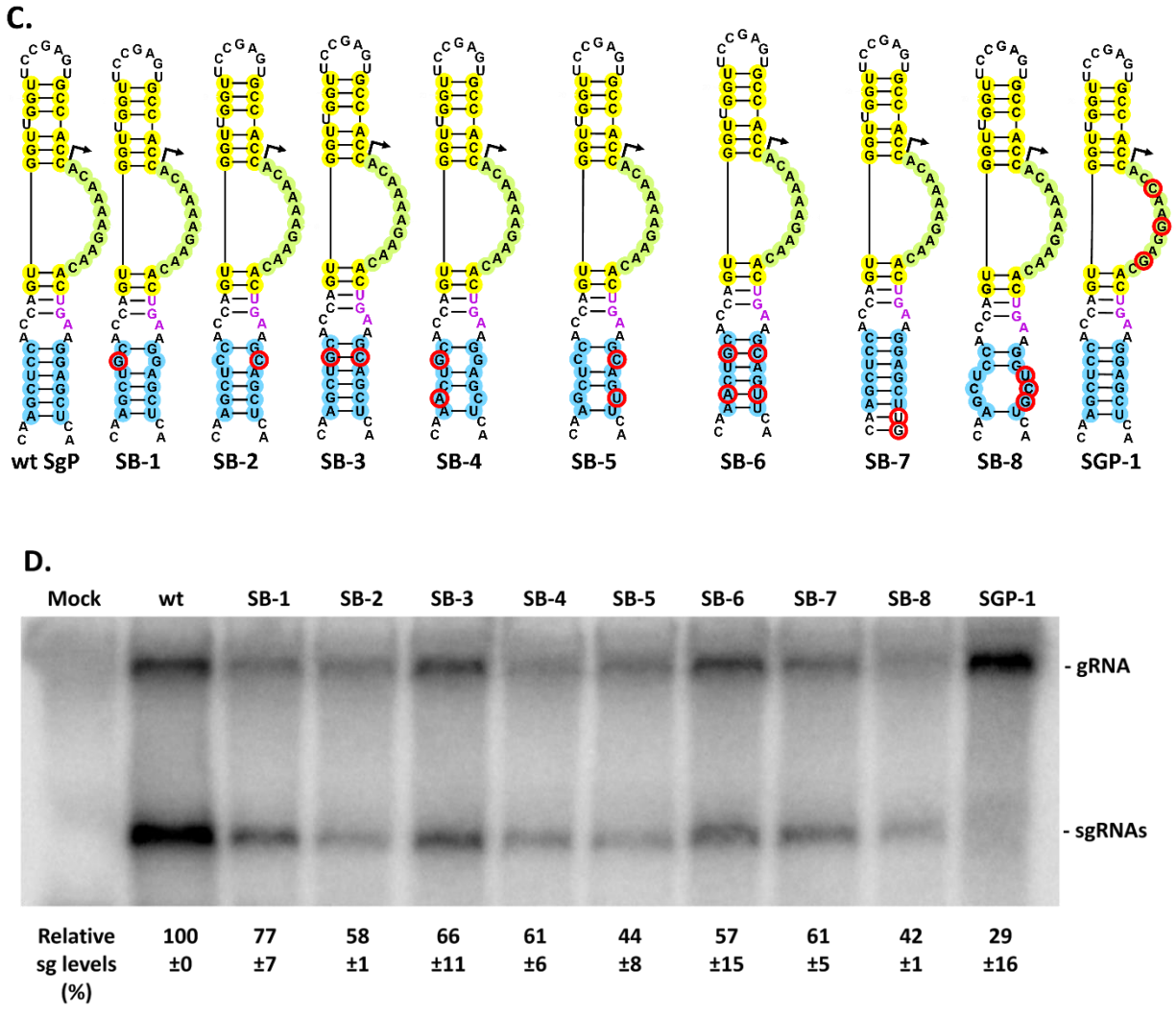


Figure 12. Analysis of subgenomic promoter mutants. A. SY subgenomic promoter mutants. Nucleotide substitutions are circled in red. **B.** Northern blot of SY mutants. **C.** SB and SGP subgenomic promoter mutants. **D.** Northern blot of SB and SGP mutants. The northern blots show viral RNA accumulation from pea protoplast infections. Percent of relative sg levels compared to wt (set to 100%) are shown with ±SE (n=4 for SY northern blot and n=3 for SB northern blot).

3.2 Investigating svRNA production and accumulation

3.2.1 Objective 1: Demonstrating xrRNA-dependent generation of svRNA *in vitro* using yeast Xrn1

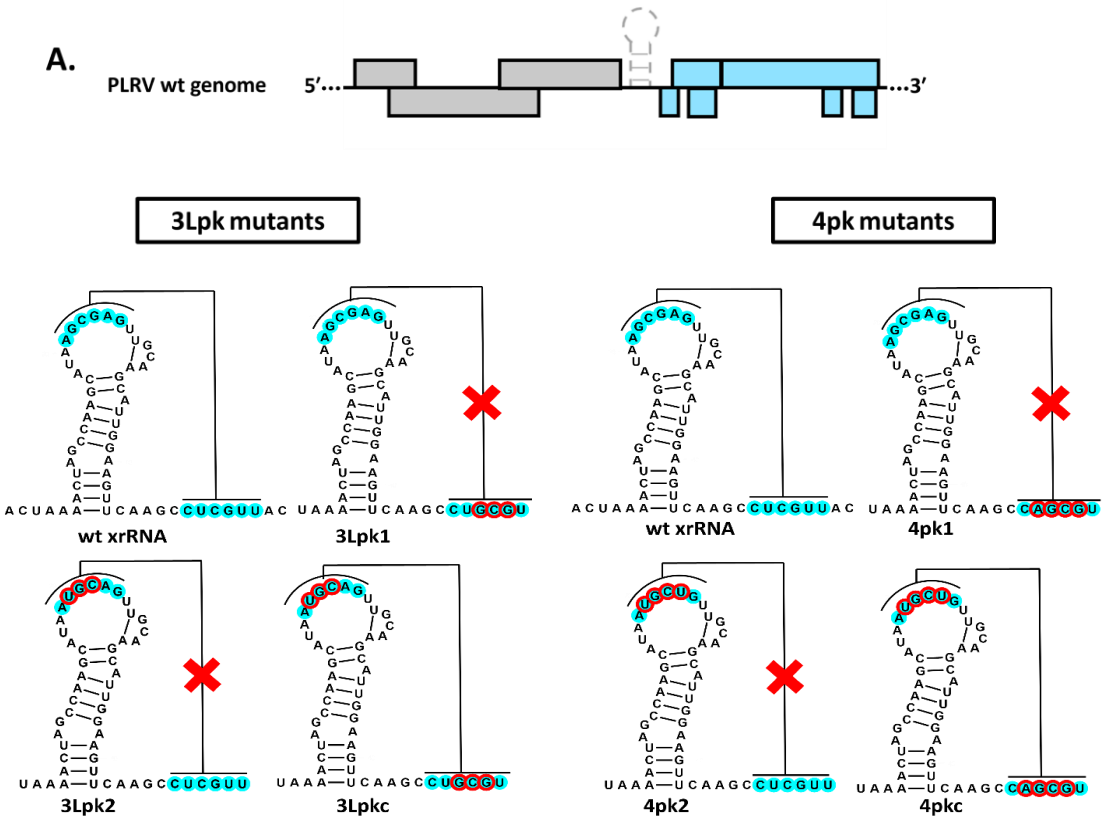
Another strain of PLRV was shown to contain an Xrn-stalling xrRNA structure that functioned through the formation of a pseudoknot (pk) (Steckelberg et al., 2018, 2020). This was demonstrated using a short fragment containing the xrRNA by *in vitro* Xrn degradation assays. A similar pk-containing xrRNA is also predicted just downstream of the sg mRNA initiation site, in our strain of PLRV (gene bank: KP090166.1; **Figure 13A**). To test if this xrRNA was active, two sets of compensatory mutants were constructed. The 3Lpk set involved substitutions in 3 out of the 6 nts involved in the pk interaction, and 4pk set involved substitutions in 4 out of the 6 nts involved in the pk interaction (**Figure 13A**). 3Lpk1 and 4pk1 were predicted to destabilize the pk that should result in no svRNA formation. Similarly, 3Lpk2 and 4pk2 were predicted to destabilize the pk. In contrast, restorative mutants 3Lpkc and 4pkc were designed to restore the pk and svRNA accumulation (**Figure 13A**).

Yeast Xrn1 degradation assay with wt mono-phosphorylated PLRV genome showed that Xrn1 generates an svRNA fragment that comigrates with the 2481 nt long svRNA size marker transcript in agarose gels (**Figure 13B**). All negative controls (reactions without Xrn1) failed to produce any svRNA (**Figure 13B**). Mutants 3Lpk1, 3Lpk2, 4pk1 and 4pk2 reactions with Xrn1 were not able to generate svRNA fragments, while in mutants 3Lpkc and 4pkc accumulation of the svRNA was restored (**Figure 13B**). The *in vitro* assays indicate that the xrRNA structure is active within the full-length PLRV genome, however the levels of accumulation of svRNA are quite low, suggesting that the xrRNA has weak stalling activity in this RNA context and assay.

3.2.2 Objective 2: Determining if svRNA accumulates in pea protoplast infections

The same genomic mutants containing compensatory mutations in the pk described in the previous section were tested for svRNA production in pea protoplast infections. Unfortunately, the 26 nt difference in length between the svRNA and the sg mRNA meant that the two RNAs would comigrate and not be resolved by gel electrophoresis. However, a reduction in the combined sg mRNA/svRNA band for xrRNA inactivating mutants would suggest that svRNA was being produced. The northern blot for the 3Lpk set shows increased accumulation of the combined band for mutants 3Lpk1, 3Lpk2 and 3Lpkc (**Figure 13C**). The

greatest increase was seen for 3Lpk1 with 120% accumulation of sgRNAs. The northern blot for the 4pk set shows a modest decrease in the combined band to 90% and 91%, respectively, for 4pk1 and 4pk2 (**Figure 13C**). There was also an increase in sgRNAs accumulation for the compensatory 4pkc (125% sgRNAs accumulation). This latter result with the 4pk mutant set is consistent with production of a small amount of svRNA, however because the 3Lpk mutant set gave negative results, a clear conclusion cannot be made.



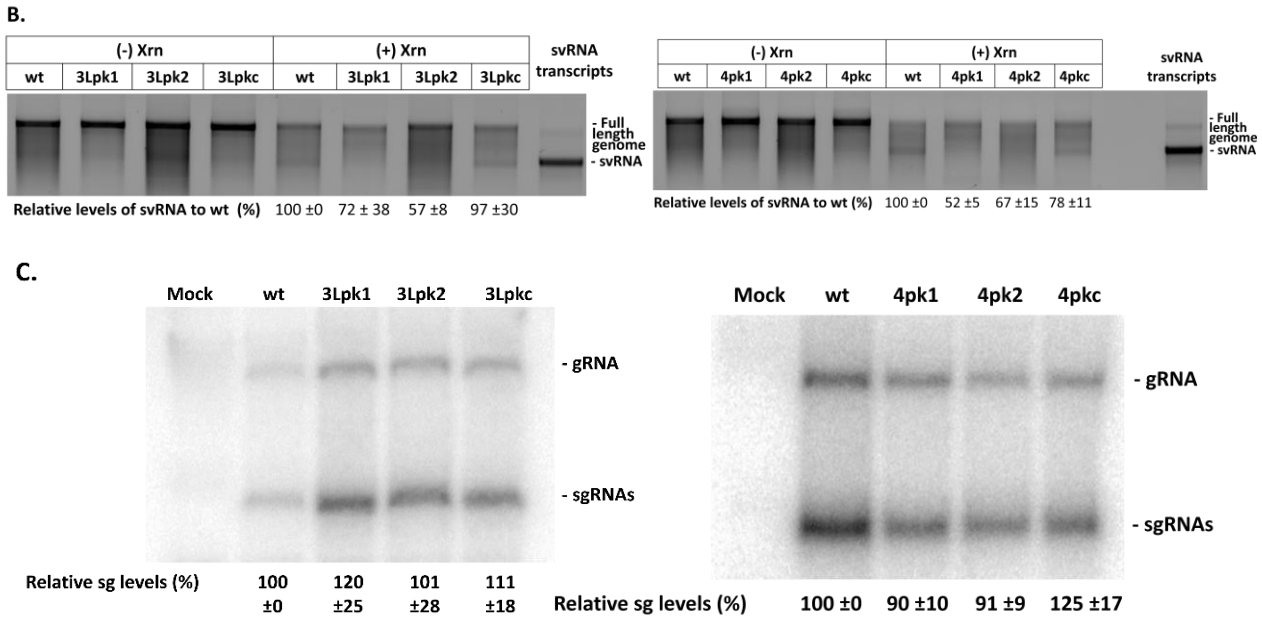


Figure 13. xrRNA compensatory PLRV genomic mutants and corresponding yeast Xrn1 *in vitro* degradation assays and northern blots. A. xrRNA compensatory mutant 3Lpk and 4pk sets. The location of the xrRNA in the PLRV genome is indicated by the light grey dashed line, outline of a stem loop present in the intergenic region. The nucleotides that are involved in the pseudoknot (pk) interaction are highlighted in light blue. Modified nucleotides in each mutant are outlined in red. Black lines without red "X" corresponds to pk formation, while black lines with a red "X" depicts the pk's inability to form. **B.** Yeast Xrn1 *in vitro* degradation assays of 3Lpk and 4pk mutant sets. Both assays were conducted at 37°C for 30 mins. RNA fragments were separated in a 2% agarose gel stained with ethidium bromide. svRNA transcript marker was used in both gels to identify the degradation product. Percent relative levels of remaining genome and relative levels of svRNA to wt (set to 100%) are shown with ±SE (n=3). **C.** Northern blots of 3Lpk and 4pk mutant sets. The northern blots show pea protoplast infections, that involved transfecting 10 µg of capped RNA transcripts into 1,000,000 protoplasts. Infections were then incubated for 24 hours at 22°C and total nucleic acids were extracted and run in a 2% agarose gel at 150 V for 45 minutes. The agarose gel was transferred onto a positive charged nylon membrane and blotted with several 3'-terminal oligo probes labeled with γ^{32}P. Blots were exposed to a phosphorimager for 1 week before visualization and quantification using Quantity One®. The accumulation of viral genomic RNA is labeled as gRNA and accumulation of sg mRNA and svRNA is labeled as sgRNAs. Percent of relative sg levels compared to wt (set to 100%) are shown with ±SE (n=4).

3.3 Investigating protein translation from sg mRNA and svRNA

3.3.1 Objective 1: Characterization of sg mRNA translation products *in vitro*

It was shown in previous work that the PLRV sg mRNA coding capacity consists of CP, MP and CP-RTD (Chkuaseli & White, 2022; Tacke et al., 1990; Xu et al., 2018). Due to the presence of a p3a ORF in the sg mRNA, it was also proposed that the PLRV sg mRNA is capable of expressing a 5 kDa p3a product, along with the protein products listed above (Smirnova et al., 2015). However, p3a would be translated from a non-AUG start codon, AUA.

In order to characterize the protein expression from the sg mRNA, a wheat germ extract (wge) *in vitro* translation system was used in which protein products were labelled with ^{35}S-methionine (^{35}S-Met). The protein products of each translation reaction were then separated in SDS-PAGE and visualized by exposure to a phosphorimager. Capped and uncapped genome and sg mRNA transcripts were initially tested. The genome and sg mRNA normally contain a VPg covalently linked to their 5'-ends (**Figure 14A**), and when present it could facilitate translation. However, it is not possible to add the VPg to *in vitro* transcripts, so instead, 5'-capped versions were tested to promote maximal translation from the sg mRNA. The mock reaction was the negative control, where no viral RNA transcripts were added. The results in **Figure 14B** show the proteins translated from the wt genome and wt sg mRNA. Both uncapped and capped genome transcripts were capable of producing P1/2 (RdRp, 139 kDa), P1 (70 kDa), and P0 (28 kDa) (**Figure 14B, grey highlights**). Uncapped and capped sg mRNA transcripts produced CP-RTD (99 kDa), CP (23 kDa), MP (17 kDa) and p3a (5 kDa) (**Figure 14B, green highlights**). Quantification of the proteins in **Figure 14B** shows that capped transcripts generated more protein products than uncapped transcripts; likely due to the cap structure providing enhanced translation initiation and mRNA stability. As predicted, the sg mRNA encodes CP, CP-RTD, and MP, as well as p3a, even though it has a non-AUG start codon.

3.3.2 Objective 2: Comparing sg mRNA and svRNA translation products *in vitro*

Earlier results (**Figure 13**) indicated that svRNA can be generated *in vitro* and may accumulate to low levels during infections. The svRNA is 26 nts shorter than the sg mRNA and contains the same encoded proteins. Therefore, svRNAs ability to act as a message for the translation of these proteins was tested.

Translation of svRNA was tested and compared to sg mRNA translation. The results in **Figure 14C** shows that svRNA capped and uncapped transcripts were capable of producing the same protein products as sg mRNA capped and uncapped transcripts (*i.e.,* CP, MP and p3a). A slight increase in CP accumulation levels were seen when svRNA capped transcripts were compared to sg mRNA capped transcripts, while the opposite was observed for the uncapped counterparts. These finding indicate that uncapped svRNAs (as predicted to exist in infections) have the ability to permit a low level of translation of its encoded proteins and therefore could potentially act as a viral message during infections.

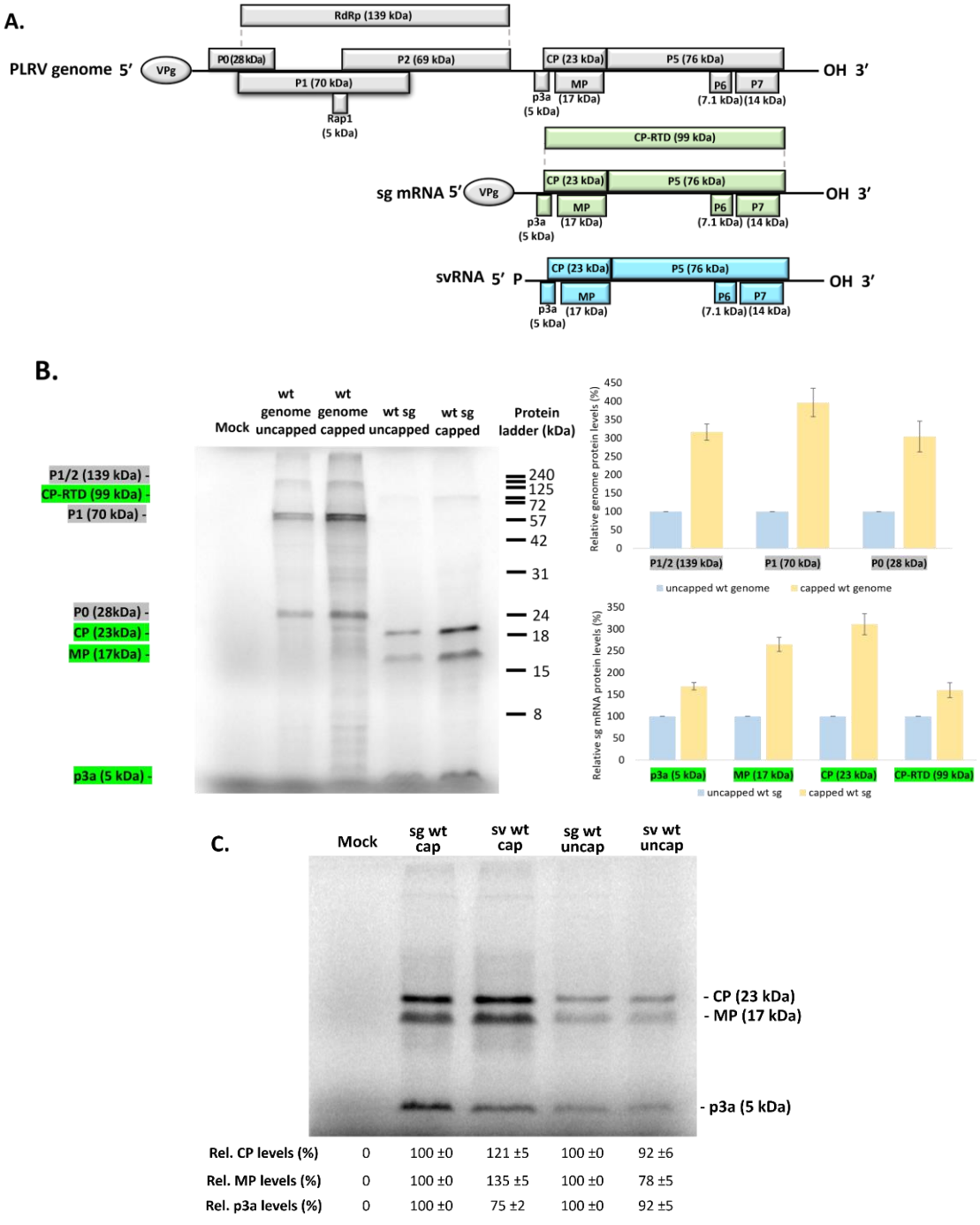


Figure 14. *In vitro* translation of PLRV genome, sg mRNA and svRNA. A. PLRV genome, sg mRNA and svRNA outline. Depicting encoded proteins corresponding to translated proteins in part B and C. **B.** Capped versus uncapped genome and sg mRNA translation. Wheat germ extract system was used, along with labeling proteins with ^{35}S-Met to visualize translation products. BLUeye prestained protein ladder (FroggaBio) was used to verify the identity of each protein. The gel was run for 55 mins at 200 V, then exposed onto a phosphorimager for 1 week before visualization using Quantity One®. Bar graphs depict percent of relative protein accumulation levels compared to uncapped wt genome and uncapped sg (set at 100%), ±SE presented as error bars (n=3). **C.** Comparison of sg mRNA and svRNA translation. Percent of relative protein accumulation levels compared to uncapped and capped sg (set at 100%) are shown, along with ±SE (n=3).

3.3.3 Objective 3: Determining if a downstream stem loop (dSL) enhances p3a production

The p3a protein is encoded in the sg mRNA and svRNA, and a potential dSL was identified, by comparative sequence analysis, 17 nts downstream from the p3a non-canonical AUA start codon (**Figure 15A**). RNA secondary structures of various polerovirus dSLs were generated using the *RNAStructure* folding software (Reuter et al., 2010) (**Figure 15B**). The conservation and positioning of the dSLs suggest that they could enhance translation from the non-AUG start codons by stalling the scanning 43S ribosome subunit over the initiation site. In order to test this hypothesis, the sg mRNA and svRNA of PLRV were analyzed.

However, before investigating the possible involvement of the dSL in p3a production, the identity of the p3a product generated by *in vitro* translation was confirmed. To do this, the AUA start codon was replaced by AAA in a p3a KO mutant and with AUG in a p3a AUG mutant (**Figure 16A, B**). These changes were made in both sg mRNA and svRNA. The results from the translation assays showed that p3a non-AUG start codon KO mutants reduced p3a accumulation, while optimal p3a AUG mutants enhanced p3a levels (**Figure 16C**), confirming the identity of the p3a band.

Having confirmed the identity of p3a (5 kDa) in the system, dSL mutants in sg mRNA and svRNA were generated and tested. The first three dSL mutants (m1, m2, m3) were designed to destabilize the dSL through wobble base pair mutations (**Figure 17A**). The other three dSL mutants were a compensatory mutant set that would destabilize the dSL (mutants KO1, KO2) and re-establish the dSL (mutant C3) (**Figure 17A**). Mutants were tested for both sg mRNA and svRNA and the results indicated primarily minor effects on p3a levels (**Figure 17B, C**), suggesting that the dSL does not notably contribute to translation of p3a.

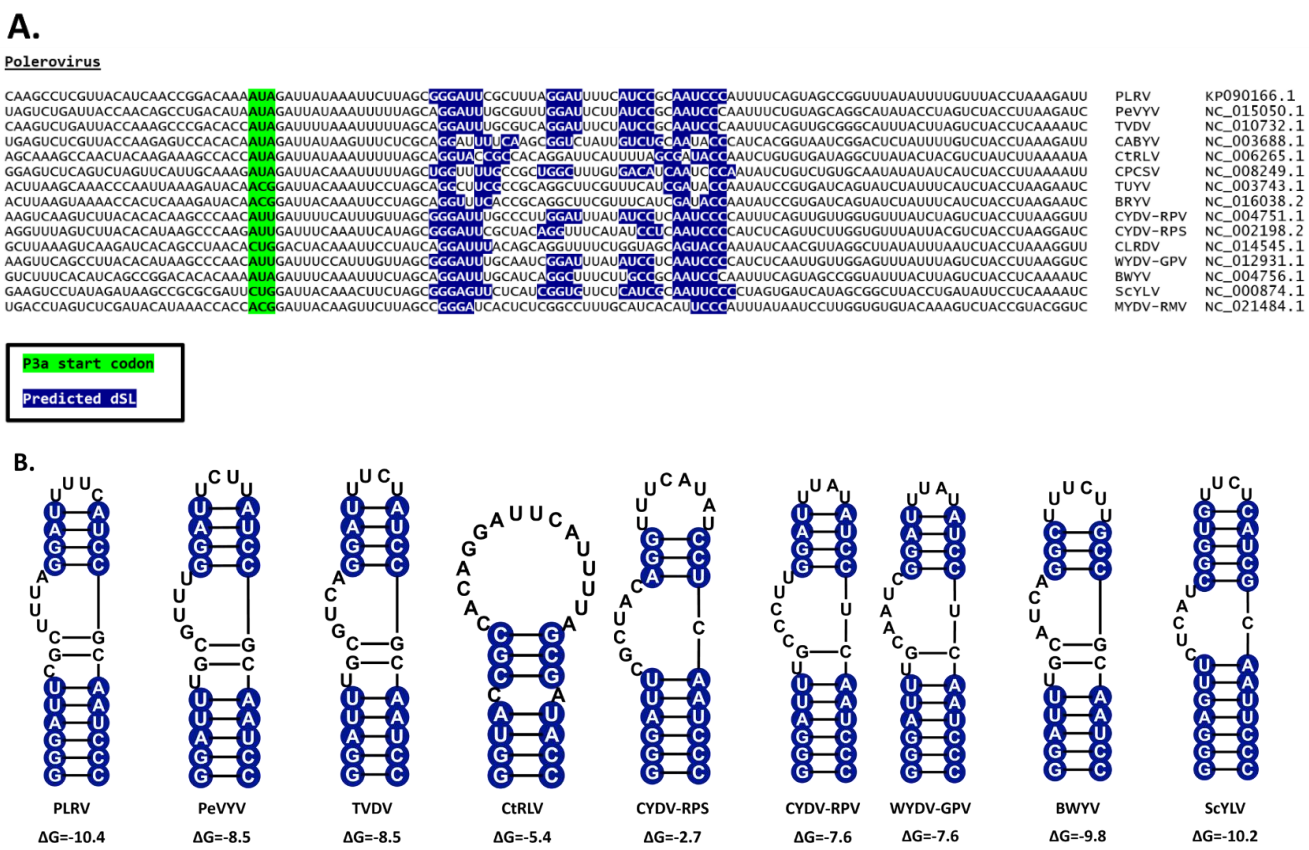
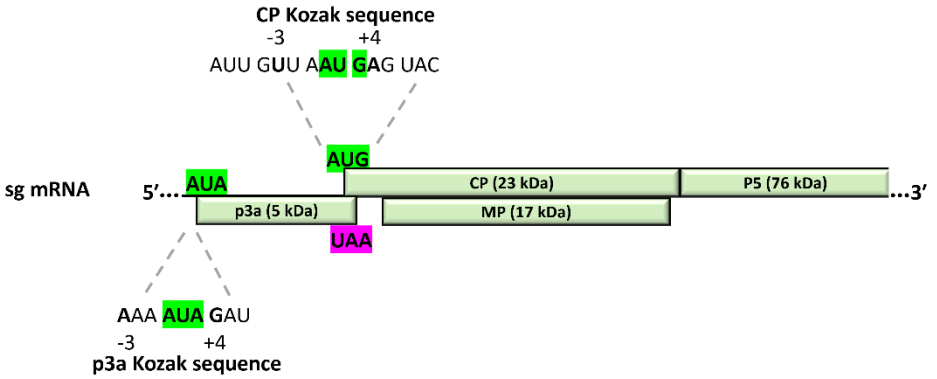


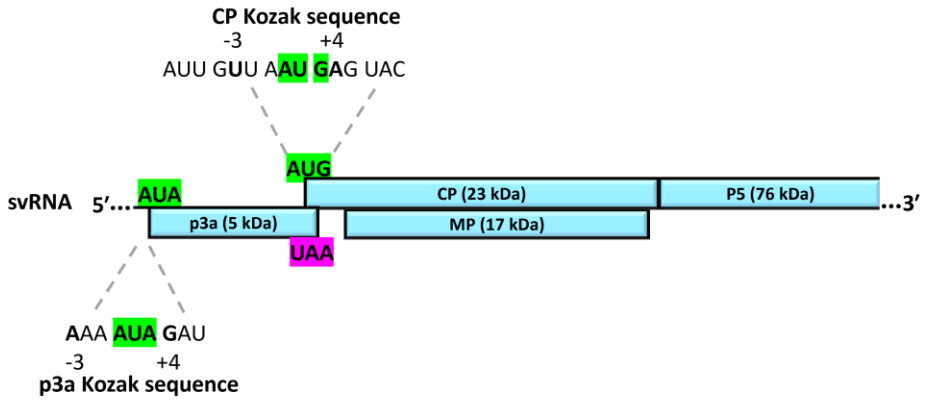
Figure 15. Comparative analysis of potential downstream stem loop (dSL) sequences in poleroviruses. **A.** Polerovirus sequences and NCBI accession numbers are from Smirnova et al., 2015. Dark blue highlighted nucleotides written in white indicate the potential conserved dSL. Green highlighted nucleotides depict the p3a non-AUG start codon. **B.** *RNAStructure* fold predictions of poleroviruses that form a dSL equivalent to PLRV. Dark blue highlighted nucleotides, written in white correspond to the nucleotides shown in part A. Folding predictions are shown for Pepper vein yellows virus (PeVYV, NC_015050.1), Tobacco vein distorting virus (TVDV, NC_010732.1), Carrot red leaf virus (CtRLV, NC_006265.1), Cereal yellow dwarf virus-RPS (CYDV-RPS, NC_002198.2), Cereal yellow dwarf virus-RPV (CYDV-RPV, NC_004751.1), Wheat yellow dwarf virus-GPV (WYDV-GPV, NC_012931.1), Beet western yellows virus (BWYV, NC_004756.1), and Sugarcane yellow leaf virus (ScYLV, NC_000874.1). Corresponding ΔG for each local fold are written below each structure.

A.



wt sg mRNA	92 - **A**AA **AUA** **G**AU UAU AAA UUC ..... 203 - UCA AUU G**U**U A**AU** **G**AG UAC GGU CGU GGU **UAA**
sg p3a KO	92 - **A**AA A**A**A **G**AU UAU AAA UUC ..... 203 - UCA AUU G**U**U A**AU** **G**AG UAC GGU CGU GGU **UAA**
sg p3a AUG	92 - **A**AA **AU**G **G**AU UAU AAA UUC ..... 203 - UCA AUU G**U**U A**AU** **G**AG UAC GGU CGU GGU **UAA**

B.



wt svRNA	66 - **A**AA **AUA** **G**AU UAU AAA UUC ..... 177 - UCA AUU GUU A**AU** **G**AG UAC GGU CGU GGU **UAA**
sv p3a KO	66 - **A**AA A**A**A **G**AU UAU AAA UUC ..... 177 - UCA AUU GUU A**AU** **G**AG UAC GGU CGU GGU **UAA**
sv p3a AUG	66 - **A**AA **AU**G **G**AU UAU AAA UUC ..... 177 - UCA AUU GUU A**AU** **G**AG UAC GGU CGU GGU **UAA**

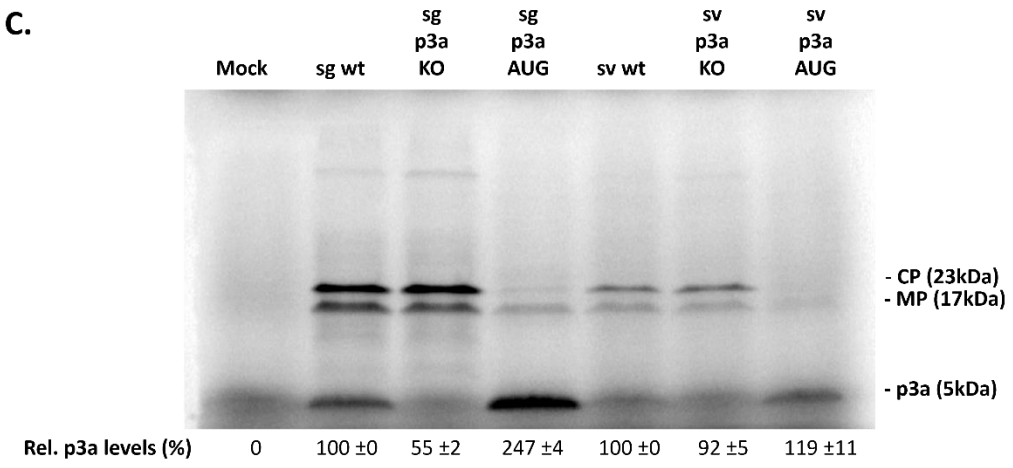


Figure 16. sg mRNA and svRNA p3a KO and p3a AUG mutant diagram and *in vitro* translation results. A. and **B.** show modifications made in the p3a start codon indicated in red and underlined. Location of the p3a start codon is shown in the sg mRNA and svRNA organization above the mutants. The strong Kozak sequence of p3a and weaker Kozak sequence of CP are depicted as well. **C.** *In vitro* translation results of the mutants shown in parts A and B. sg mRNA translation reactions used capped transcripts, while svRNA reactions used uncapped transcripts. Mock reaction involved using no viral transcript. All reactions were run in SDS-PAGE (5% stacking, 12% resolving) for 30 minutes at 200 V. Then gels were dried and exposed onto a phosphorimager for 1 week before visualization and quantification with Quantity One®. Percent of relative protein accumulation compared to wt (set at 100%) are shown below the gel, along with ±SE (n=3).

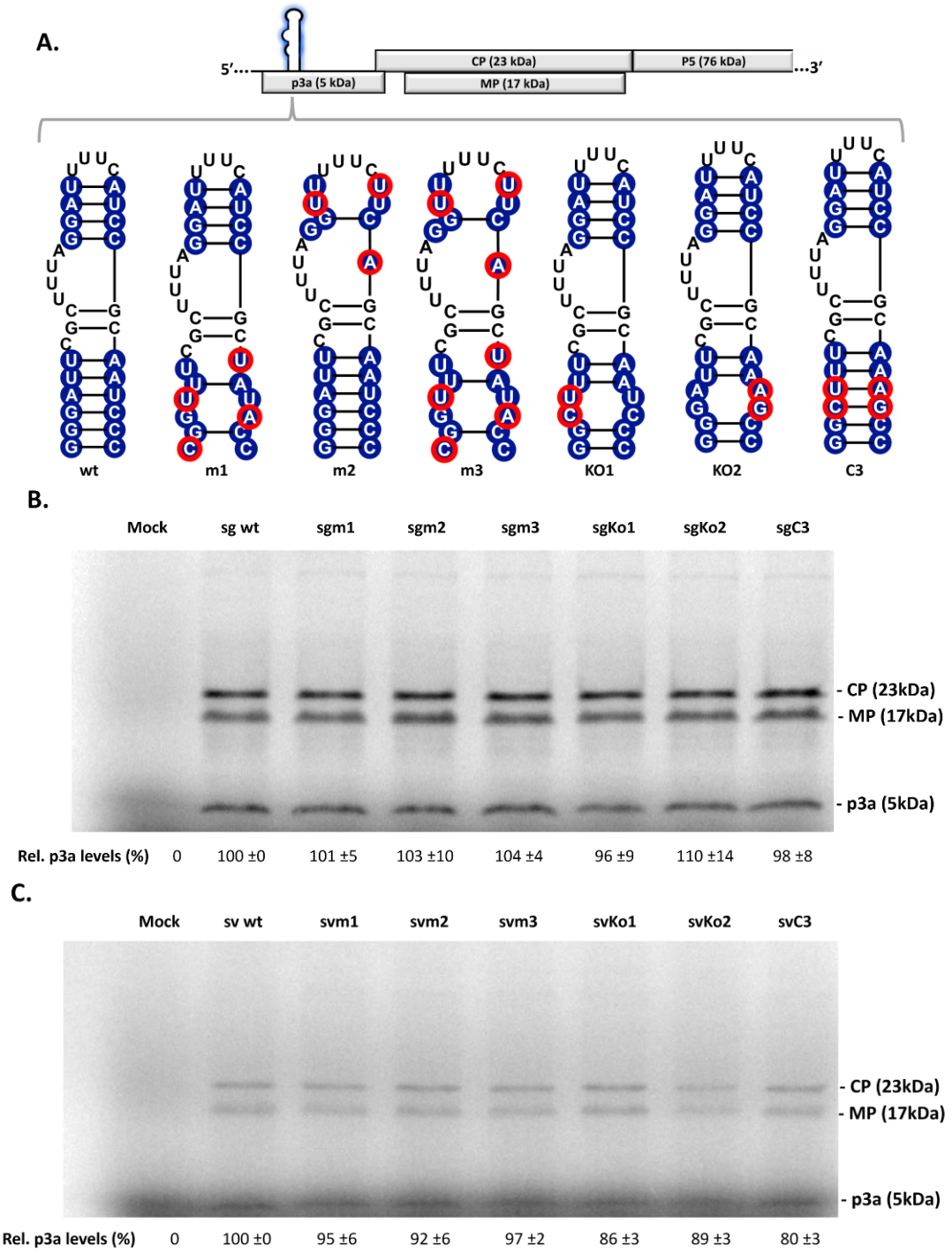


Figure 17. dSL mutants along with corresponding *in vitro* translation results. A. dSL mutants. wt dSL (stem loop structure highlighted in dark blue) is depicted in the sg mRNA organization (genome coordinates 3488 nt-3526 nt) within the p3a ORF. Modified nucleotides are outline in red. Dark blue highlighted nucleotides correspond to highlighted nucleotides in figure 15. **B.** and **C.** *In vitro* translation results of dSL inserted in sg mRNA and svRNA. Both SDS-PAGEs show *in vitro* translation experiments in wheat germ extract system using capped sg mRNA transcripts (B.) and uncapped sv RNA transcripts (C.). Proteins were labeled with ^{35}S-Met for visualization and quantified with Quantity One®. Both SDS-PAGEs (5% stacking, 12% resolving) were run for 30 minutes at 200 V and exposed onto a phosphorimager for 1 week. Percent of relative protein accumulation compared to wt (set to 100%) are shown, with ±SE (n=3).

4 Discussion

4.1 Summary of key findings

This investigation of PLRV's gene expression strategies revealed several interesting findings. First, analysis of the SgP region identified an RNA secondary structure and a linear sequence that are important for efficient sg mRNA transcription. Second, the putative xrRNA structure was shown to be moderately active *in vitro* in the full-length genome, however the accumulation of the svRNA during infections could not be confirmed. Third, both the sg mRNA and svRNA were able to act as messages for translation of their encoded proteins, however the putative dSL did not act as an enhancer of p3a translation. These results are discussed in more detail below.

4.2 Exploring PLRV's gene expression strategies

4.2.1 sg mRNA transcription in PLRV

Sg mRNAs are produced during infections to allow for expression of a subset of viral genes. Viruses use different mechanisms to transcribe their sg mRNAs. Determining the mechanism used requires detailed analysis of the SgP region where sg mRNA transcription initiates. In PLRV, there was a problem, because the SgP is located in the RdRp coding region. One approach to deal with this is to move a copy of the promoter region to a different non-essential region of the viral genome. This method was used previously to study the SgP in turnip crinkle virus (TCV) (Wang & Simon, 1997; Wu et al., 2010). However, this strategy did not work with the PLRV SgP (**Figure 11B**). The PLRV genome SgP insertion mutants could have failed due to the SgP sequence not folding correctly in their new locations. Another possibility is that the SgP segment that was moved was missing some important flanking sequences or structures. This last possibility could be tested by inserting larger segments of the SgP region. Since the transplanted SgP mutants did not work, an alternative approach of introducing silent substitutions into the SgP in its natural RdRp-coding region was used (**Figure 12A and C**).

Sequence and structural analysis of the SgP identified a conserved sequence (green) and two possible secondary structures (yellow and blue) (**Figure 10**). The green sequence that corresponds to the 5′-end of the sg mRNA was important and is part of the SgP (**Figure 12D**). Analysis of the blue interaction indicated that it was not required for efficient sg mRNA

transcription (**Figure 12D**). The lack of importance in the blue interacting nucleotides is consistent with this interaction not being as absolutely conserved throughout poleroviruses, in contrast to the yellow interaction (**Figure 10**). When the yellow interaction was tested, it was found to be important for efficient sg mRNA accumulation (**Figure 12B**). Therefore, like the green sequence, the yellow interaction can be considered part of the SgP.

The yellow interaction could potentially occur in either the plus- or minus-strand of the genome (**Figure 10**). If it functions in the plus-stand, then it could act as an attenuation structure to help stall the RdRp during minus-strand synthesis. This role would correspond with transcription occurring by a premature termination mechanism (**Figure 3C**), with the attenuation structure formed by a local interaction. In contrast, viruses in the *Tombusviridae* family, such as tomato bushy stunt virus (TBSV), consist of attenuation structures with complex interactions that are reliant on multiple long-distance interactions or, as in red clover necrotic mosaic virus (RCNMV), an attenuation structure that forms between two genomic RNAs (Sit et al., 1998; White & Nagy, 2004). Currently, only plant viruses in the family *Tombusviridae* are known to use the premature termination mechanism (Jiwan & White, 2011; Newburn & White, 2015), and poleroviruses are not related to tombusvirids. Therefore, if a premature termination mechanism does occur in PLRV, it would be the first example outside of the family *Tombusviridae*.

In contrast, if the yellow interaction occurs in the minus-stand, then it would be consistent with an internal initiation mechanism of sg mRNA transcription (**Figure 3A**). The role of start site-adjacent RNA secondary structure in internal initiation mechanism is to bind the RdRp, as has been shown for brome mosaic virus, BMV (Adkins & Kao, 1998; Haasnoot et al., 2000, 2002). However, the observation that opposes this mechanism are minus-strand secondary structure predictions (**Figure 10B**). The minus-strand secondary structure predictions for the yellow interaction in the SgP region show little conservation, suggesting that the minus-strand structure likely has no significant function. In contrast, the more consistent formation of the yellow interaction in the plus-strand tends to support a premature termination mechanism. The importance of the green sequence and its similarity to the 5′-terminus of the genome is also in agreement with a premature termination mechanism, because these two sequences

show very high levels of identity in other viruses that use this mechanism (Jiwan & White, 2011).

This initial SgP analysis has provided information on what is important for sg mRNA transcription in PLRV, *i.e.*, the green sequence and yellow interaction. The current data seem more consistent with a premature termination mechanism, however additional studies will be needed to help determine the actual mechanism being employed by PLRV.

4.2.2 PLRV's small viral RNA (svRNA)

The presence of an xrRNA structure within a viral genome leads to the production of svRNA by preventing 5′-to-3′ exoribonuclease from degrading the entire viral RNA (**Figure 8**) (Vicens & Kieft, 2021). The svRNAs generated can be either non-coding or coding (Steckelberg et al., 2018). The first xrRNA element and corresponding svRNA were identified in a dianthovirus, RCNMV (Iwakawa et al., 2008). This particular xrRNA was located in the 3′-untranslated region (3′-UTR) of the viral genome and resulted in accumulation of non-coding RNA degradation product termed SR1f. The non-coding svRNA consisted of translational elements, suggesting its potential involvement in translation regulation (Iwakawa et al., 2008). It was recently discovered that xrRNAs could also be located internally within viral RNA genomes, potentially resulting in the production of coding RNA decay intermediates that act as mRNAs (Steckelberg et al., 2018). PLRV's xrRNA, which is located internally within the viral RNA genome, is predicted to generate an svRNA that encodes several viral proteins (**Figure 5, light blue ORFs**) (Steckelberg et al., 2018). *In vitro* Xrn1 degradation assays confirmed this svRNA could be generated from the PLRV genome (**Figure 13B**).

Since PLRV can potentially produce an svRNA that encodes viral proteins, it was proposed that it could function as an additional viral mRNA during infections. There is an existing example of svRNAs functioning as mRNA in opium poppy mosaic virus (OPMV), a tombusvirid (Ilyas et al., 2021). OPMV has an xrRNA that is located upstream of the known sg mRNA transcription start site and thus produces an svRNA that is slightly longer than the sg mRNA. *In vitro* translation results showed that the OPMV svRNA encodes an additional protein that is not expressed from the sg mRNA (Ilyas et al., 2021). In contrast, PLRV's svRNA is smaller than the sg mRNA, and *in vitro* translation results showed that svRNA produces the same proteins as sg mRNA (**Figure 14C**); which seems redundant. The results also suggested that, if

the PLRV svRNA is produced during infections, it would be at very low levels, because (i) disrupting the essential pk in the xrRNA structure had little effect on the band intensity corresponding to sg mRNA and svRNA (**Figure 13C**) and (ii) inactivating the SgP to eliminate sg mRNA accumulation did not reveal much of a residual band corresponding to svRNA (**Figure 12B and D**). This suggests that very little svRNA is present during infections and, although it could potentially allow for translation of some viral proteins, its contribution would be quite small.

Currently, the function of the svRNA in PLRV is not clear. The fact that an active xrRNA structure is maintained in the viral genome suggests that the svRNA serves some function that provides a selective advantage. As mentioned above, production of the svRNA could boost the amount of certain 3′-encoded proteins, but only minimally. It is possible that small increases in these proteins could be advantageous in certain situations. Another possibility is that the svRNA degradation product could have preceded the sg mRNA as the sole message for translating ORFs at the 3′-end of the viral genome. However, over time the development of the SgP and sg mRNA could have begun the slow process of replacing the svRNA translational template with sg mRNA. In this scenario, the svRNA would eventually become dispensable and be fully replaced by the sg mRNA. This seems possible, as not all poleroviruses contain xrRNAs and svRNA, suggesting that they are not essential (Steckelberg et al., 2018).

4.3 PLRV sg mRNA and svRNA translation products

It is unknown whether poleroviruses sg mRNAs have a VPg at its 5′-end, or whether the VPg acts as a translational enhancer, as in potyviruses (Duprat et al., 2002; Jiang & Laliberté, 2011). When assessing the protein products of the sg mRNA and svRNA in the *in vitro* translation system, both capped and uncapped transcripts were tested. If the sg mRNA has a VPg and it acts as a translation enhancer, then the capped version of the sg mRNA studied would be closest to the wt situation. However, if the sg mRNA has no VPg, then the uncapped sg mRNA analyzed would be closer to the wt situation. Since the 5′-end of the svRNA is produced by an exoribonuclease, it would not contain a VPg. As, predicted, lower levels of the viral proteins (p3a, CP, MP and CP-RTD) were produced from the uncapped versions of sg mRNA and svRNA (**Figure 14C, 16C, 17B, 17C**). Importantly, the results confirmed that svRNA could potentially act as a mRNA for these proteins, but very inefficiently from its uncapped message

(**Figure 14C**). The p3a protein translated from sg mRNA was of particular interest because it initiates with a non-AUG start codon, AUA. This protein was also shown to be produced in *in vitro* translation assays of other poleroviruses such as turnip yellows virus (TuYV) (Smirnova et al., 2015). The same study also demonstrated that p3a is required for systemic movement of the infection within plants, however a detailed assessment of p3a translation was not performed.

4.3.1 The dSL and p3a translation

Non-AUG start codons are usually inefficiently recognized by scanning ribosomes, but can be made to be more efficient by a strong initiation context and/or a downstream RNA secondary structure that promotes ribosome pausing on the non-AUG initiation site (**Figure 2G**) (Firth & Brierley, 2012; Geng et al., 2021; Kozak, 1990). This mechanism was previously seen in mammalian cells, however viruses such as flaviviruses and alphaviruses use downstream stem loops to enhance translation of a suboptimal initiation context (Clyde & Harris, 2006; Firth & Brierley, 2012; Ventoso et al., 2006).

In poleroviruses, a conserved dSL structure was predicted about 17 nts away from the p3a start codon by comparative structural analysis (**Figure 15**). The spacing of the dSLs would allow scanning ribosomal subunits to pause over the p3a non-AUG start codons, thereby enhancing initiation of translation. However, when testing compensatory and destabilizing dSL mutants in PLRV's sg mRNA using *in vitro* translation, p3a accumulation remained relatively constant at near wt levels (**Figure 17B, C**). This indicates that the dSL does not influence PLRV p3a translation *in vitro*. Instead, the adenylate at -3 and guanylate at +4 corresponding to an optimal Kozak consensus appears to be enough to allow for sufficient p3a translation. It is possible that *in vivo* the dSL may help to enhance translation, but this would need to be experimentally tested.

5 Future Directions

5.1 Further analysis on PLRV RNA Secondary structures

Many in lab experiments use selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) or in-line probing to analysis RNA base pairing regulatory structures found in various tombusvirids (Chkuaseli & White, 2020, 2023; Gunawardene et al., 2019; Im et al., 2021, 2023). SHAPE was also recently used to study a readthrough element in PLRV (Chkuaseli &

White, 2022), however it was not yet used to analysis RNA secondary structures such as SgP or dSL in PLRV. Using both SHAPE and in-line probing would help to confirm the *RNAStructure* predictions of SgP and dSL (**Figure 10, Figure 15**). SHAPE consists of modification of unconstrained nucleotides, followed by fluorescently labeled primer extension analysis and capillary electrophoresis to determine SHAPE reactivities (Low & Weeks, 2010). In-line probing involves nuclease cleavage between two neighboring nucleotides that are aligned in a specific conformation around the phosphate backbone, followed by γ-^{32}P labeling and separation of cleaved fragments by gel electrophoresis (Nahvi & Green, 2013). Both methods function to identify flexible and presumably unpaired nucleotides and this information is incorporated into *RNAStructure* predicting programs to generate more accurate RNA secondary structure model (Low & Weeks, 2010). Having more reliable RNA secondary structure models for both SgP and dSL would assist in investigating their functions.

5.2 Further analysis on svRNA identification

In the experiments that have been completed throughout this thesis sg mRNA KO mutants and svRNA KO mutants have been tested separately in pea protoplast infections and visualized using northern blots. Northern blots showing results of sg mRNA KO mutants and svRNA KO mutants, always show a remaining sgRNAs band regardless of the sg mRNA being KO or the svRNA being KO. It could be assumed that the remaining sgRNAs band on the sg mRNA KO mutant northern blot (**Figure 12B**) is the svRNA and vice a versa for the svRNA KO mutant northern blot (**Figure 13C**), however this was never proven. To verify the remaining sgRNAs band on northern blots additional mutants would have to be created that KO both the sg mRNA and svRNA. If these mutants result in knocking out the full and complete sgRNAs band, then the above assumption about the remaining sgRNAs band would be correct.

5.3 Further analysis on the dSL

In previous published work on RCNMV and PEMV1, western blotting was used along with *in vitro* translation and northern blotting to assess RNA secondary structures that were involved in translation. In both experiments the corresponding results from northern blotting, western blotting, and *in vitro* translation were consistent (Chkuaseli & White, 2022; Im et al., 2023). Throughout this thesis the dSL mutants were only tested using *in vitro* translation in a

wheat germ extract system (**Figure 17**), but not in virus infections. In the future, the dSL mutants consisting of a hemagglutinin (HA) tagged p3a could be transfected into pea protoplast and p3a could be detected through western blotting using an anti-HA-peroxidase high affinity (3F10) rat monoclonal antibody (Roche). If the results are consistent with the *in vitro* translation results (**Figure 17**), then this would confirm that the dSL does not contribute to p3a translation.

6 Conclusion

This study has provided new information on gene expression in PLRV. First, regulatory sequences and structures important for transcription of the sg mRNA were identified. Second, *in vitro* assays confirmed that an internal xrRNA structure was active and its svRNA product acts as a message for viral proteins, however its accumulation and function during infections could not be confirmed. Third, a predicted dSL was shown not to be involved in enhancing p3a translation from sg mRNA and svRNA. Together these findings have advanced our understanding of gene expression in PLRV.

7 Reference

Acheson, N. H. (2011). *Fundamentals of Molecular Virology* (2nd edition). Wiley.

Adkins, S., & Kao, C. C. (1998). Subgenomic RNA Promoters Dictate the Mode of Recognition by Bromoviral RNA-Dependent RNA Polymerases. *Virology*, *252*(1), 1–8. https://doi.org/10.1006/viro.1998.9449

Aitken, C. E., & Lorsch, J. R. (2012). A mechanistic overview of translation initiation in eukaryotes. *Nature Structural & Molecular Biology*, *19*(6), 568–576. https://doi.org/10.1038/nsmb.2303

Baumberger, N., Tsai, C.-H., Lie, M., Havecker, E., & Baulcombe, D. C. (2007). The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. *Current Biology: CB*, *17*(18), 1609–1614. https://doi.org/10.1016/j.cub.2007.08.039

Bortolamiol, D., Pazhouhandeh, M., Marrocco, K., Genschik, P., & Ziegler-Graff, V. (2007). The Polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. *Current Biology: CB*, *17*(18), 1615–1621. https://doi.org/10.1016/j.cub.2007.07.061

Bruyère, A., Brault, V., Ziegler-Graff, V., Simonis, M.-T., Van Den Heuvel, J. F. J. M., Richards, K., Guilley, H., Jonard, G., & Herrbach, E. (1997). Effects of Mutations in the Beet Western Yellows Virus Readthrough Protein on Its Expression and Packaging and on Virus Accumulation, Symptoms, and Aphid Transmission. *Virology*, *230*(2), 323–334. https://doi.org/10.1006/viro.1997.8476

Chkuaseli, T., & White, K. A. (2020). Activation of viral transcription by stepwise largescale folding of an RNA virus genome. *Nucleic Acids Research*, *48*(16), 9285–9300. https://doi.org/10.1093/nar/gkaa675

Chkuaseli, T., & White, K. A. (2022). Complex and simple translational readthrough signals in pea enation mosaic virus 1 and potato leafroll virus, respectively. *PLOS Pathogens*, *18*(9), e1010888. https://doi.org/10.1371/journal.ppat.1010888

Chkuaseli, T., & White, K. A. (2023). Dimerization of an umbravirus RNA genome activates subgenomic mRNA transcription. *Nucleic Acids Research*, *51*(16), 8787–8804. https://doi.org/10.1093/nar/gkad550

Cimino, P. A., Nicholson, B. L., Wu, B., Xu, W., & White, K. A. (2011). Multifaceted Regulation of Translational Readthrough by RNA Replication Elements in a Tombusvirus. *PLOS Pathogens*, *7*(12), e1002423. https://doi.org/10.1371/journal.ppat.1002423

Clyde, K., & Harris, E. (2006). RNA Secondary Structure in the Coding Region of Dengue Virus Type 2 Directs Translation Start Codon Selection and Is Required for Viral Replication. *Journal of Virology*, *80*(5), 2170–2182. https://doi.org/10.1128/jvi.80.5.2170-2182.2006

Danthinne, X., Seurinck, J., Meulewaeter, F., Van Montagu, M., & Cornelissen, M. (1993). The 3' untranslated region of satellite tobacco necrosis virus RNA stimulates translation in vitro. *Molecular and Cellular Biology*, *13*(6), 3340–3349.

DeBlasio, S. L., Xu, Y., Johnson, R. S., Rebelo, A. R., MacCoss, M. J., Gray, S. M., & Heck, M. (2018). The Interaction Dynamics of Two Potato Leafroll Virus Movement Proteins Affects Their Localization to the Outer Membranes of Mitochondria and Plastids. *Viruses*, *10*(11), 585. https://doi.org/10.3390/v10110585

Delfosse, V. C., Barón, M. P. B., & Distéfano, A. J. (2021). What we know about poleroviruses: Advances in understanding the functions of polerovirus proteins. *Plant Pathology*, *70*(5), 1047–1061. https://doi.org/10.1111/ppa.13368

Dever, T. E., & Green, R. (2012). The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, *4*(7), a013706. https://doi.org/10.1101/cshperspect.a013706

Dietzgen, R. G., Mann, K. S., & Johnson, K. N. (2016). Plant Virus–Insect Vector Interactions: Current and Potential Future Research Directions. *Viruses*, *8*(11), 303. https://doi.org/10.3390/v8110303

Dimitrov, D. S. (2004). Virus entry: Molecular mechanisms and biomedical applications. *Nature Reviews Microbiology*, *2*(2), Article 2. https://doi.org/10.1038/nrmicro817

Dinesh-Kumar, S. P., & Miller, W. A. (1993). Control of start codon choice on a plant viral RNA encoding overlapping genes. *The Plant Cell*, *5*(6), 679–692.

Dorokhov, Yu. L., Ivanov, P. A., Komarova, T. V., Skulachev, M. V., & Atabekov, J. G. (2006). An internal ribosome entry site located upstream of the crucifer-infecting tobamovirus coat protein (CP) gene can be used for CP synthesis in vivo. *Journal of General Virology*, *87*(9), 2693–2697. https://doi.org/10.1099/vir.0.82095-0

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K. S., & Robaglia, C. (2002). The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *The Plant Journal*, *32*(6), 927–934. https://doi.org/10.1046/j.1365-313X.2002.01481.x

Firth, A. E., & Brierley, I. 2012. (2012). Non-canonical translation in RNA viruses. *Journal of General Virology*, *93*(7), 1385–1409. https://doi.org/10.1099/vir.0.042499-0

French, R., & Ahlquist, P. (1988). Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. *Journal of Virology*, *62*(7), 2411–2420. https://doi.org/10.1128/jvi.62.7.2411-2420.1988

Fuller, F., Bhown, A. S., & Bishop, D. H. (1983). Bunyavirus nucleoprotein, N, and a non-structural protein, NSS, are coded by overlapping reading frames in the S RNA. *The Journal of General Virology*, *64 (Pt 8)*, 1705–1714. https://doi.org/10.1099/0022-1317-64-8-1705

Gao, F., & Simon, A. E. (2016). Multiple Cis-acting elements modulate programmed -1 ribosomal frameshifting in Pea enation mosaic virus. *Nucleic Acids Research*, *44*(2), 878–895. https://doi.org/10.1093/nar/gkv1241

Geng, G., Wang, D., Liu, Z., Wang, Y., Zhu, M., Cao, X., Yu, C., & Yuan, X. (2021). Translation of Plant RNA Viruses. *Viruses*, *13*(12), 2499. https://doi.org/10.3390/v13122499

Gill, C. C., & Chong, J. (1975). Development of the infection in oat leaves inoculated with barley yellow dwarf virus. *Virology*, *66*(2), 440–453. https://doi.org/10.1016/0042-6822(75)90216-0

Gunawardene, C. D., Newburn, L. R., & White, K. A. (2019). A 212-nt long RNA structure in the Tobacco necrosis virus-D RNA genome is resistant to Xrn degradation. *Nucleic Acids Research*, *47*(17), 9329–9342. https://doi.org/10.1093/nar/gkz668

Haasnoot, P. C., Brederode, F. T., Olsthoorn, R. C., & Bol, J. F. (2000). A conserved hairpin structure in Alfamovirus and Bromovirus subgenomic promoters is required for efficient RNA synthesis in vitro. *RNA*, *6*(5), 708–716.

Haasnoot, P. C. J., Olsthoorn, R. C. L., & Bol, J. F. (2002). The Brome mosaic virus subgenomic promoter hairpin is structurally similar to the iron-responsive element and functionally equivalent to the minus-strand core promoter stem-loop C. *RNA (New York, N.Y.)*, *8*(1), 110–122. https://doi.org/10.1017/s1355838202012074

Hoang, H.-D., Neault, S., Pelin, A., & Alain, T. (2021). Emerging translation strategies during virus-host interaction. *Wiley Interdisciplinary Reviews. RNA*, *12*(1), e1619. https://doi.org/10.1002/wrna.1619

Hull, R. (2001). *Matthews' Plant Virology, Fourth Edition* (4th edition). Academic Press.

Ilyas, M., Du, Z., & Simon, A. E. (2021). Opium Poppy Mosaic Virus Has an Xrn-Resistant, Translated Subgenomic RNA and a BTE 3' CITE. *Journal of Virology*, *95*(9), e02109-20. https://doi.org/10.1128/JVI.02109-20

Im, J. S. H., Newburn, L. R., Kent, G., & White, K. A. (2021). Trans-Activator Binding Site Context in RCNMV Modulates Subgenomic mRNA Transcription. *Viruses*, *13*(11), Article 11. https://doi.org/10.3390/v13112252

Im, J. S. H., Sheppard, J. R., & White, K. A. (2023). Structure-Based Regulatory Role for the 5'UTR of RCNMV RNA2. *Viruses*, *15*(3), Article 3. https://doi.org/10.3390/v15030722

Iwakawa, H., Mizumoto, H., Nagano, H., Imoto, Y., Takigawa, K., Sarawaneeyaruk, S., Kaido, M., Mise, K., & Okuno, T. (2008). A Viral Noncoding RNA Generated by cis-Element-Mediated Protection against 5'→3' RNA Decay Represses both Cap-Independent and Cap-Dependent Translation. *Journal of Virology*, *82*(20), 10162–10174. https://doi.org/10.1128/JVI.01027-08

Jaag, H. M., Kawchuk, L., Rohde, W., Fischer, R., Emans, N., & Prüfer, D. (2003). An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein. *Proceedings of the National Academy of Sciences*, *100*(15), 8939–8944. https://doi.org/10.1073/pnas.1332697100

Jacks, T., Madhani, H. D., Masiarz, F. R., & Varmus, H. E. (1988). Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell*, *55*(3), 447–458. https://doi.org/10.1016/0092-8674(88)90031-1

Jang, C. J., Lo, M. C. Y., & Jan, E. (2009). Conserved element of the dicistrovirus IGR IRES that mimics an E-site tRNA/ribosome interaction mediates multiple functions. *Journal of Molecular Biology*, *387*(1), 42–58. https://doi.org/10.1016/j.jmb.2009.01.042

Jiang, J., & Laliberté, J.-F. (2011). The genome-linked protein VPg of plant viruses-a protein with many partners. *Current Opinion in Virology*, *1*(5), 347–354. https://doi.org/10.1016/j.coviro.2011.09.010

Jiwan, S. D., & White, K. A. (2011). Subgenomic mRNA transcription in Tombusviridae. *RNA Biology*, *8*(2), 287–294. https://doi.org/10.4161/rna.8.2.15195

Jones, C. I., Zabolotskaya, M. V., & Newbury, S. F. (2012). The 5' → 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development. *Wiley Interdisciplinary Reviews. RNA*, *3*(4), 455–468. https://doi.org/10.1002/wrna.1109

Kastenmayer, J. P., & Green, P. J. (2000). Novel features of the XRN-family in Arabidopsis: Evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13985–13990.

Koev, G., & Miller, W. A. (2000). A Positive-Strand RNA Virus with Three Very Different Subgenomic RNA Promoters. *Journal of Virology*, *74*(13), 5988–5996.

Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., Zerbini, F. M., & Kuhn, J. H. (2020). Global Organization and Proposed Megataxonomy of the Virus World. *Microbiology and Molecular Biology Reviews: MMBR*, *84*(2), e00061-19. https://doi.org/10.1128/MMBR.00061-19

Koonin, E. V., Krupovic, M., & Agol, V. I. (2021). The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiology and Molecular Biology Reviews : MMBR*, *85*(3), e00053-21. https://doi.org/10.1128/MMBR.00053-21

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Research*, *9*(20), 5233–5252. https://doi.org/10.1093/nar/9.20.5233

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, *44*(2), 283–292. https://doi.org/10.1016/0092-8674(86)90762-2

Kozak, M. (1989). The scanning model for translation: An update. *The Journal of Cell Biology*, *108*(2), 229–241. https://doi.org/10.1083/jcb.108.2.229

Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(21), 8301–8305. https://doi.org/10.1073/pnas.87.21.8301

Kozieł, E., Otulak-Kozieł, K., & Bujarski, J. J. (2021). Plant Cell Wall as a Key Player During Resistant and Susceptible Plant-Virus Interactions. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.656809

Kumar, D., Kumar, R., Hyun, T. K., & Kim, J.-Y. (2015). Cell-to-cell movement of viruses via plasmodesmata. *Journal of Plant Research*, *128*(1), 37–47. https://doi.org/10.1007/s10265-014-0683-6

Low, J. T., & Weeks, K. M. (2010). SHAPE-directed RNA secondary structure prediction. *Methods (San Diego, Calif.)*, *52*(2), 150–158. https://doi.org/10.1016/j.ymeth.2010.06.007

Meulewaeter, F., Danthinne, X., Va, M., Montagu, & Cornelissen, M. (1998). 5′- and 3′-sequences of satellite tobacco necrosis virus RNA promoting translation in tobacco. *The Plant Journal*, *14*(2), 169–176. https://doi.org/10.1046/j.1365-313X.1998.00104.x

Michaeli, S., Clavel, M., Lechner, E., Viotti, C., Wu, J., Dubois, M., Hacquard, T., Derrien, B., Izquierdo, E., Lecorbeiller, M., Bouteiller, N., De Cilia, J., Ziegler-Graff, V., Vaucheret, H., Galili, G., & Genschik, P. (2019). The viral F-box protein P0 induces an ER-derived autophagy degradation pathway for the clearance of membrane-bound AGO1. *Proceedings of the National Academy of Sciences*, *116*(45), 22872–22883. https://doi.org/10.1073/pnas.1912222116

Miller, J. S., & Mayo, M. A. (1991). The location of the 5′ end of the potato leafroll luteovirus subgenomic coat protein mRNA. *The Journal of General Virology*, *72 (Pt 11)*, 2633–2638. https://doi.org/10.1099/0022-1317-72-11-2633

Miras, M., Miller, W. A., Truniger, V., & Aranda, M. A. (2017). Non-canonical Translation in Plant RNA Viruses. *Frontiers in Plant Science*, *8*. https://www.frontiersin.org/articles/10.3389/fpls.2017.00494

Mugridge, J. S., Coller, J., & Gross, J. D. (2018). Structural and molecular mechanisms for the control of eukaryotic 5'-3' mRNA decay. *Nature Structural & Molecular Biology*, *25*(12), 1077–1085. https://doi.org/10.1038/s41594-018-0164-z

Nagarajan, V. K., Jones, C. I., Newbury, S. F., & Green, P. J. (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(6), 590–603. https://doi.org/10.1016/j.bbagrm.2013.03.005

Nagy, P. D., Strating, J. R. P. M., & van Kuppeveld, F. J. M. (2016). Building Viral Replication Organelles: Close Encounters of the Membrane Types. *PLoS Pathogens*, *12*(10), e1005912. https://doi.org/10.1371/journal.ppat.1005912

Nahvi, A., & Green, R. (2013). Chapter Twenty-Two—Structural Analysis of RNA Backbone Using In-Line Probing. In J. Lorsch (Ed.), *Methods in Enzymology* (Vol. 530, pp. 381–397). Academic Press. https://doi.org/10.1016/B978-0-12-420037-1.00022-1

Newburn, L. R., Nicholson, B. L., Yosefi, M., Cimino, P. A., & White, K. A. (2014). Translational readthrough in Tobacco necrosis virus-D. *Virology*, *450–451*, 258–265. https://doi.org/10.1016/j.virol.2013.12.006

Newburn, L. R., & White, K. A. (2015). Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology*, *479–480*, 434–443. https://doi.org/10.1016/j.virol.2015.02.032

Newburn, L. R., & White, K. A. (2020). A trans-activator-like structure in RCNMV RNA1 evokes the origin of the trans-activator in RNA2. *PLOS Pathogens*, *16*(1), e1008271. https://doi.org/10.1371/journal.ppat.1008271

Nicholson, B. L., & White, K. A. (2011). 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. *Current Opinion in Virology*, *1*(5), 373–380. https://doi.org/10.1016/j.coviro.2011.10.002

Nicholson, B. L., Wu, B., Chevtchenko, I., & White, K. A. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA*, *16*(7), 1402–1419. https://doi.org/10.1261/rna.2135210

Pazhouhandeh, M., Dieterle, M., Marrocco, K., Lechner, E., Berry, B., Brault, V., Hemmer, O., Kretsch, T., Richards, K. E., Genschik, P., & Ziegler-Graff, V. (2006). F-box-like domain in the polerovirus protein P0 is required for silencing suppressor function. *Proceedings of the National Academy of Sciences*, *103*(6), 1994–1999. https://doi.org/10.1073/pnas.0510784103

Pelletier, J., & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, *334*(6180), Article 6180. https://doi.org/10.1038/334320a0

Peter, K. A., Liang, D., Palukaitis, P., & Gray, S. M. (2008). Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *Journal of General Virology*, *89*(8), 2037–2045. https://doi.org/10.1099/vir.0.83625-0

https://doi.org/10.1016/j.chom.2008.10.007

Prüfer, D., Kawchuk, L., Monecke, M., Nowok, S., Fischer, R., & Rohde, W. (1999). Immunological analysis of potato leafroll luteovirus (PLRV) P1 expression identifies a 25 kDa RNA-binding protein derived via P1 processing. *Nucleic Acids Research*, *27*(2), 421–425. https://doi.org/10.1093/nar/27.2.421

Rampersad, S., & Tennant, P. (2018). Replication and Expression Strategies of Viruses. *Viruses*, 55–82. https://doi.org/10.1016/B978-0-12-811257-1.00003-6

Reuter, J.S., & Mathews, D. H. (2010). RNAstructure: software for RNA secondary structure predictions and analysis. BMC Bioinformativs. 11,129.

Riegger, R. J., & Caliskan, N. (2022). Thinking Outside the Frame: Impacting Genomes Capacity by Programmed Ribosomal Frameshifting. *Frontiers in Molecular Biosciences*, *9*. https://www.frontiersin.org/articles/10.3389/fmolb.2022.842261

Rodnina, M. V., Korniy, N., Klimova, M., Karki, P., Peng, B.-Z., Senyushkina, T., Belardinelli, R., Maracci, C., Wohlgemuth, I., Samatova, E., & Peske, F. (2020). Translational recoding: Canonical translation mechanisms reinterpreted. *Nucleic Acids Research*, *48*(3), 1056–1067. https://doi.org/10.1093/nar/gkz783

Ryabova, L. A., Pooggin, M. M., & Hohn, T. (2006). Translation reinitiation and leaky scanning in plant viruses. *Virus Research*, *119*(1), 52–62. https://doi.org/10.1016/j.virusres.2005.10.017

Sevvana, M., Klose, T., & Rossmann, M. G. (2021). Principles of Virus Structure. *Encyclopedia of Virology*, 257–277. https://doi.org/10.1016/B978-0-12-814515-9.00033-3

Shi, X., Zhang, Z., Zhang, C., Zhou, X., Zhang, D., & Liu, Y. (2021). The molecular mechanism of efficient transmission of plant viruses in variable virus–vector–plant interactions. *Horticultural Plant Journal*, *7*(6), 501–508. https://doi.org/10.1016/j.hpj.2021.04.006

Shivprasad, S., Pogue, G. P., Lewandowski, D. J., Hidalgo, J., Donson, J., Grill, L. K., & Dawson, W. O. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. *Virology*, *255*(2), 312–323. https://doi.org/10.1006/viro.1998.9579

Siegel, R. W., Adkins, S., & Kao, C. C. (1997). Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proceedings of the National Academy of Sciences*, *94*(21), 11238–11243. https://doi.org/10.1073/pnas.94.21.11238

Sit, T. L., Vaewhongs, A. A., & Lommel, S. A. (1998). RNA-Mediated Trans-Activation of Transcription from a Viral RNA. *Science*, *281*(5378), 829–832. https://doi.org/10.1126/science.281.5378.829

Smirnova, E., Firth, A. E., Miller, W. A., Scheidecker, D., Brault, V., Reinbold, C., Rakotondrafara, A. M., Chung, B. Y.-W., & Ziegler-Graff, V. (2015). Discovery of a Small Non-AUG-Initiated ORF in Poleroviruses and Luteoviruses That Is Required for Long-Distance Movement. *PLOS Pathogens*, *11*(5), e1004868. https://doi.org/10.1371/journal.ppat.1004868

Sola, I., Almazán, F., Zúñiga, S., & Enjuanes, L. (2015). Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annual Review of Virology*, *2*(1), 265–288. https://doi.org/10.1146/annurev-virology-100114-055218

Spall, V. E., Shanks, M., & Lomonossoff, G. P. (1997). Polyprotein Processing as a Strategy for Gene Expression in RNA Viruses. *Seminars in Virology*, *8*(1), 15–23. https://doi.org/10.1006/smvy.1997.0102

Steckelberg, A.-L., Vicens, Q., Costantino, D. A., Nix, J. C., & Kieft, J. S. (2020). The crystal structure of a Polerovirus exoribonuclease-resistant RNA shows how diverse sequences are integrated into a conserved fold. *RNA*, *26*(12). https://doi.org/10.1261/rna.076224.120

Steckelberg, A.-L., Vicens, Q., & Kieft, J. S. (2018). Exoribonuclease-Resistant RNAs Exist within both Coding and Noncoding Subgenomic RNAs. *mBio*, *9*(6), e02461-18. https://doi.org/10.1128/mBio.02461-18

Sztuba-Solińska, J., Stollar, V., & Bujarski, J. J. (2011). Subgenomic messenger RNAs: Mastering regulation of (+)-strand RNA virus life cycle. *Virology*, *412*(2), 245–255. https://doi.org/10.1016/j.virol.2011.02.007

Tacke, E., Prüfer, D., Salamini, F., & Rohde, W. (1990). Characterization of a potato leafroll luteovirus subgenomic RNA: Differential expression by internal translation initiation and UAG suppression. *The Journal of General Virology*, *71 (Pt 10)*, 2265–2272. https://doi.org/10.1099/0022-1317-71-10-2265

Taliansky, M., Mayo, M. A., & Barker, H. (2003). Potato leafroll virus: A classic pathogen shows some new tricks. *Molecular Plant Pathology*, *4*(2), 81–89. https://doi.org/10.1046/j.1364-3703.2003.00153.x

Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A.-C., & Vagner, S. (2003). Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biology of the Cell*, *95*(3–4), 169–178. https://doi.org/10.1016/S0248-4900(03)00033-9

Treder, K., Kneller, E. L. P., Allen, E. M., Wang, Z., Browning, K. S., & Miller, W. A. (2008). The 3′ cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. *RNA*, *14*(1), 134–147. https://doi.org/10.1261/rna.777308

Turina, M., Desvoyes, B., & Scholthof, K.-B. G. (2000). A Gene Cluster Encoded by Panicum Mosaic Virus Is Associated with Virus Movement. *Virology*, *266*(1), 120–128. https://doi.org/10.1006/viro.1999.0069

van der Wilk, F., Verbeek, M., Dullemans, A. M., & van den Heuvel, J. F. J. M. (1997). The Genome-Linked Protein of Potato Leafroll Virus Is Located Downstream of the Putative Protease Domain of the ORF1 Product. *Virology*, *234*(2), 300–303. https://doi.org/10.1006/viro.1997.8654

Ventoso, I., Sanz, M. A., Molina, S., Berlanga, J. J., Carrasco, L., & Esteban, M. (2006). Translational resistance of late alphavirus mRNA to eIF2α phosphorylation: A strategy to overcome the antiviral effect of protein kinase PKR. *Genes & Development*, *20*(1), 87–100. https://doi.org/10.1101/gad.357006

Vicens, Q., & Kieft, J. S. (2021). Shared properties and singularities of exoribonuclease-resistant RNAs in viruses. *Computational and Structural Biotechnology Journal*, *19*, 4373–4380. https://doi.org/10.1016/j.csbj.2021.07.024

Wang, J., & Simon, A. E. (1997). Analysis of the two subgenomic RNA promoters for turnip crinkle virus in vivo and in vitro. *Virology*, *232*(1), 174–186. https://doi.org/10.1006/viro.1997.8550

Wang, Z., Kraft, J., Hui, A., & Miller, W. A. (2010). Structural plasticity of Barley yellow dwarf virus-like cap-independent translation elements in four genera of plant viral RNAs. *Virology*, *402*(1), 177–186. https://doi.org/10.1016/j.virol.2010.03.025

Wells, S. E., Hillner, P. E., Vale, R. D., & Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Molecular Cell*, *2*(1), 135–140. https://doi.org/10.1016/s1097-2765(00)80122-7

White, K. A. (2002). The premature termination model: A possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. *Virology*, *304*(2), 147–154. https://doi.org/10.1006/viro.2002.1732

White, K. A., & Nagy, P. D. (2004). Advances in the Molecular Biology of Tombusviruses: Gene Expression, Genome Replication, and Recombination. In *Progress in Nucleic Acid Research and Molecular Biology* (Vol. 78, pp. 187–226). Academic Press. https://doi.org/10.1016/S0079-6603(04)78005-8

White, K. A. (2011). Tombusvirus. In C. Tidona & G. Darai (Eds.), *The Springer Index of Viruses* (pp. 1917–1926). Springer. https://doi.org/10.1007/978-0-387-95919-1_314

Wong, J. H., Alfatah, M., Sin, M. F., Sim, H. M., Verma, C. S., Lane, D. P., & Arumugam, P. (2017). A yeast two-hybrid system for the screening and characterization of small-molecule inhibitors of protein–protein interactions identifies a novel putative Mdm2-binding site in p53. *BMC Biology*, *15*(1), 108. https://doi.org/10.1186/s12915-017-0446-7

Wu, B., Oliveri, S., Mandic, J., & White, K. A. (2010). Evidence for a Premature Termination Mechanism of Subgenomic mRNA Transcription in a Carmovirus. *Journal of Virology*, *84*(15), 7904–7907. https://doi.org/10.1128/JVI.00742-10

Xu, Y., Ju, H.-J., DeBlasio, S., Carino, E. J., Johnson, R., MacCoss, M. J., Heck, M., Miller, W. A., & Gray, S. M. (2018). A Stem-Loop Structure in Potato Leafroll Virus Open Reading Frame 5 (ORF5) Is Essential for Readthrough Translation of the Coat Protein ORF Stop Codon 700 Bases Upstream. *Journal of Virology*, *92*(11), e01544-17. https://doi.org/10.1128/JVI.01544-17

Zeenko, V., & Gallie, D. R. (2005). Cap-independent Translation of Tobacco Etch Virus Is Conferred by an RNA Pseudoknot in the 5′-Leader *. *Journal of Biological Chemistry*, *280*(29), 26813–26824. https://doi.org/10.1074/jbc.M503576200